Akita University

R/S Pox レッグライン特性に基づく ポートスキャントラフィックの検知に関する 研究

2014

髙橋　秋典

Akita University

目 次

Akita University

Akita University

図 目 次

Akita University

表 目 次

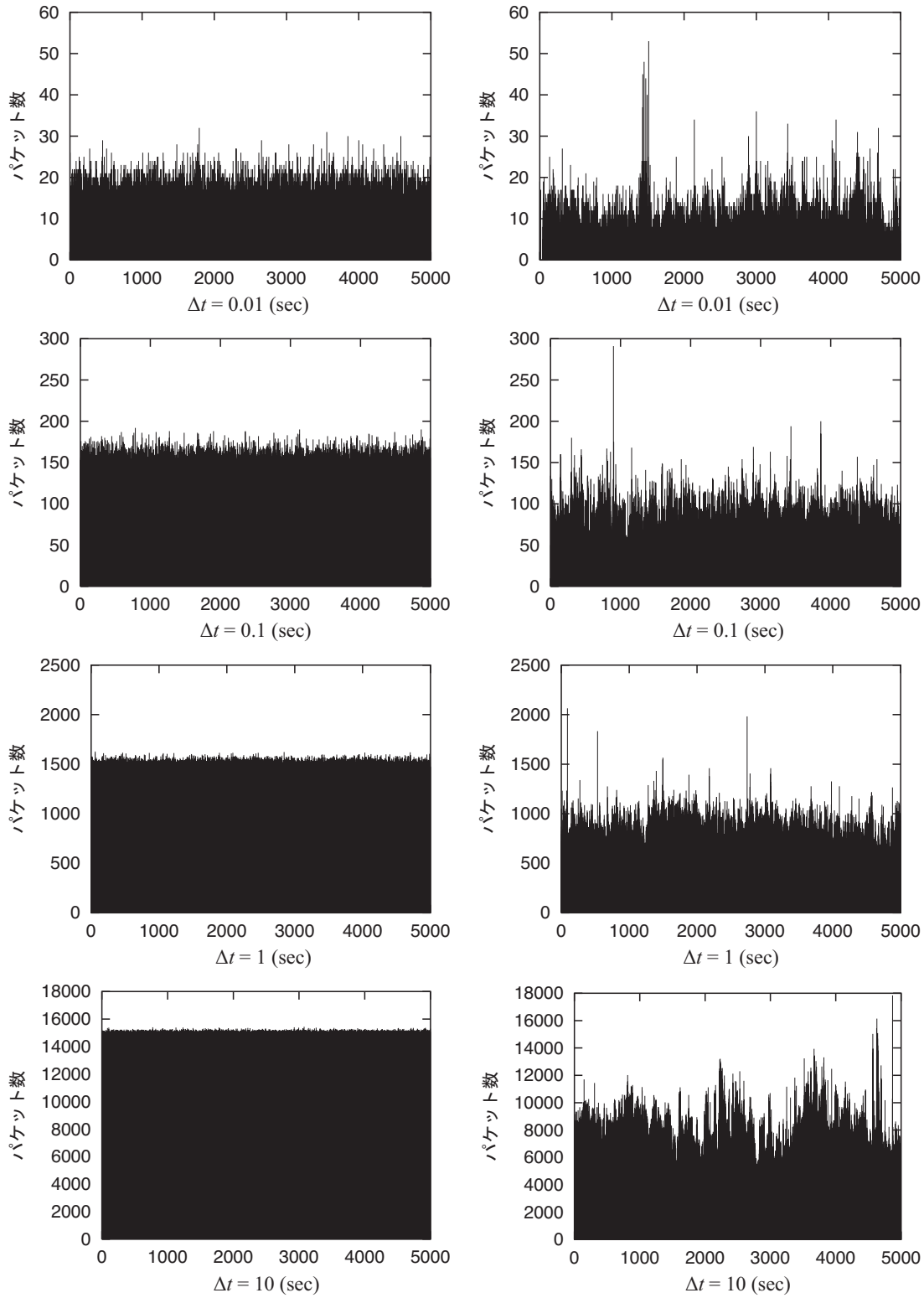
第1章　緒論

1.1　背景

近年，コンピュータの発達に伴い，インターネットの普及は著しく，電子メールやWebサービスはもちろんのこと，電子マネー，SNS，メディアコンテンツ配信など様々なサービスが運用されており，日常生活においては必要不可欠な重要なインフラの一つとなっている．このインターネット上には利便性の高い多様なアプリケーションによる情報通信が行われている一方，DDoS攻撃や事前調査を行うポートスキャン等の悪意あるトラフィックも疎通しており，これにより輻輳崩壊などによるネットワーク障害の発生が懸念される場合がある．このような問題に対処するために，トラフィックデータを解析することは，ネットワーク管理･運用，セキュリティー研究において重要な位置づけとなっている．このためトラフィック疎通における異常性の検知は，トラフィックエンジニアリングの観点からは重要性が高い．

ネットワークトラフィックの異常検知に対する現実的な方法としては，トラフィック量の異常な，或いは急激な増加を検知する手法であると考えられるが，パケットトラフィックは従来の通信路で取り扱われていたポアソン過程には従わず，自己相似性に起因する長期記憶過程である [1] ことが指摘されているため，そのトラフィック時系列X_tの解析には注意深い検討が必要である．ここで，集積時間スケールを変化させたときの実トラフィックデータから得られた時系列X_tを図1.1(a)に，および従来のトラフィックモデルであるポアソン過程に従う時系列X_tを図1.1(b)に示す [2]．それぞれの時系列は集積時間スケールΔtを0.01，0.1，1，10秒と変化させたものである．ここから，従来のトラフィックモデルにおいては集積時間スケールを大きくすると時系列が平滑化されるのに対して，自己相似性を有する実際のトラフィック時系列X_tにはバースティネスがあり，集積時間スケールを如何にあげても任意時刻で大きく変化するX_tが存在する．つまり，異常性の検知に対してトラフィック量閾値を利用する場合，正常な変化範囲のフローレベルを誤って異常と判断してしまう不都合が予想されるため，その検討には注意が必要となる．

Akita University



(a) 従来のトラフィックモデル　　(b) 実際のトラフィック

図 1.1 時間スケール変化に対するトラフィック時系列 X_t

Akita University

1.2 関連研究

インターネットのパケットトラフィック時系列の自己相似性に対しては，その要因を調査する研究が行われてきた．その要因には，上位層プロトコルのTCP輻輳制御 [3][4] や輻輳・非輻輳の臨界領域の影響 [5]，また，ネットワークアプリケーションによる影響 [6] や，DDoS攻撃のような非定常的な異常トラフィックによる影響 [7] などの報告があり，トラフィック事象変化に対して自己相似性の様相が変化するということが観測されている．

この自己相似性を表すハーストパラメータ H の導出法の一つであるR/S解析法は, H. E. Hurstがナイル川の流量変動の統計的解析に導入した後 [8], B.B. Mandelbrotにより数学的な基礎付けがなされた統計的解析法 [9] で，Leland et al.[1] が初めて，ネットワークトラフィックの自己相似性の解析に導入したものである．この解析法は ”グラフ的な方法 ”と呼ばれるもので，ハーストパラメータ H は観測時系列データにおける任意長区間から算出されるR/S統計量をグラフにしたR/S Pox Diagramのプロット点群の傾きより導出される．このR/S解析法に対しても様々な検討が行われており，R/S Pox Diagramのハーストパラメータ H 計算範囲の妥当性 [10][11] や計算高速化のためのアルゴリズム改良 [12] などが提案されている．

これらの研究においては，R/S解析法から導出されるハーストパラメータ H に着目しトラフィック事象に対する検討が行われているが，非定常的に突発的トラフィック量増加を示すレベルシフトが発生する場合や，間欠的に到着し，且つ到着期間内では周期列となるようなトラフィック時系列の場合，R/S Pox Diagramに特徴的なプロット形状が現れることが報告されている [13][14][15]．我々の調査においては，特に周期的時系列に対しては，プロット点群が一つの傾きではなく，途中から折れ曲がり，二つの傾きを呈することが観測されている．

これらのプロット形状は，非定常性を表す特徴であると考えられるが，信頼性に欠けるハーストパラメータ H 推定を与えてしまうため，統計解析においては除去すべき対象として検討されてきた．Mandelbrot et al.がコンピュータシミュレーションを用いた研究 [9] で，ランダム過程に決定論的過程，具体的にはsin波を重畳したとき，そのR/S Pox Diagram形状に変化が現れることを指摘していた．しかし，非定常性が重畳されたときのR/S Pox Diagramの形状変化を事象変化検出に積極的に利用しようとした研究は見当たらない．非定常性に対するプロット形状の定量化は，トラフィック事象変化の検知に積極的に応用できると考えられる．また，R/S解析法の計算過程において，R/S統計量を求める任意長区間のサイズによって未計算部分が存在し，直近の取得データが解析に反映されないことがある．これに対して問題視する研究も見当たらない．

Akita University

1.3 本研究の目的

本研究では，従来，トラフィック特性として解析されてきた自己相似性推定法の一つであるR/S解析法に対して検討を行い，非定常的トラフィック事象に対する特性変化を積極的に応用した新たな異常検知法を提案することを目的とする．

まず，R/S解析法の未計算区間により時間特性に対する即応性が低下する問題点を指摘し，その改良法を提案する．さらに，その改良法を用いたR/S Pox Diagramに現れる特徴的なプロット形状の要因について明らかにし，その形状を定量的に扱えるR/S Pox レッグライン特性を提案する．この名称は，周期的時系列に対して折れ曲がるようなプロット点群の形状が人間の脚部に似ていることに由来する．この特性を用いて周期的時系列に対する周期推定法を検討し，シミュレーション時系列に対する性能評価を行った．また，その際，実環境から観測された長期的ポートスキャン攻撃トラフィック [16] に対して適用を試み，そのパケット到着間隔を推定することで本手法の有効性を示した．

1.4 本論文の内容

本論文は全5章より構成され，第1章は緒論とした．

第2章では，R/S解析法による時系列解析手順に関する問題点を指摘し，効果的な改良方法について提案した．また，自己相似性を表すハーストパラメータ導出過程に生成されるR/S Pox Diagramについて，非定常性を含む時系列観測時に現れる特徴的プロット形状の要因について，シミュレーションによる調査を行い検討を加えた．その結果，R/S Pox Diagramでは，突発的トラフィック量増減を示すレベルシフト時系列や，時間間隔をおいて対象に攻撃を送信するような周期的時系列に対して，その非定常性を表現し得ることを明らかにした．特に，周期的時系列の周期について，プロット点群の折れ曲がるポイントにその特徴が明確に現れていることを明らかにした．

第3章では，第2章の検討結果に基づき，プロット形状を定量化する新たな特性として，R/S Pox レッグライン特性を提案し，その特徴量および特徴量を用いた周期推定法について定義した．この特性の名称は，R/S Pox Diagramの折れ曲がるようなプロット形状が，人体の脚部に似ていることに由来する．特性の特徴量は，ハーストパラメータの算出手順を参考に，Pox Diagramの前半部分 (ST:Slope of Thigh)，後半部分 (SS:Slope of Shin) のそれぞれ上限点群 (Sup)，平均点群 (Avg)，下限点群 (Inf) の回帰直線の傾きから算出した．また，周期を示す折れ曲がるポイントを前半部，後半部の回帰直線の交点 x 座標から推定する方法を提案した．本特

Akita University

性の有用性を検討するため，非定常性に対する特徴量の変化をシミュレーションにより検証し，第2章で推測した異常検知に対する性能を有することを確認した．また，提案した周期推定法により，ST_{Sup}を選択したときの周期的時系列の周期が推測可能であることを明らかにした．

第4章では，本特性の実トラフィック環境での異常検知性能を評価するため，実環境下で取得されたトラフィックデータを用いて，低レートのパケットで事前調査を行う長期的ポートスキャントラフィックの検知実験を行った．その結果，実際に観測されたポートスキャントラフィックの周期を，荒い精度ではあるが，推測できることを明らかにした．また，特徴量および推定された周期の経時変化を調査したところ，ポートスキャントラフィックの有無を検知できる性能を有していることを明らかにした．

第5章では，本研究で得られた主な知見と今後残された課題について述べている．

参考文献

[1] Leland, W.E., Taqqu, M.S., Willinger, W. and Wilson, D.V.: On the Self -Similar Nature of Ethernet Traffic, Computer Communications Review, Vol.23, No.4, pp.183-193 (1993).

[2] 宮林尚英：ハーストパラメータによるネットトラヒック特性の解析，平成 14 年度秋田大学大学院鉱山学研究科修士学位論文（2002）

[3] 住田義明, 大崎博之, 村田正幸, 宮原秀夫：上位層プロトコルがネットワークトラヒックの自己相似性に与える影響，電子情報通信学会論文誌 B，Vol.J82-B, No.6，pp.1126-1137（1999）.

[4] 土井博生，松田崇弘，山本幹：TCP ふくそう制御がトラヒックのマルチフラクタル性に与える影響, 電子情報通信学会論文誌 B, Vol.J88-B, No.6, pp.1029-1037（2005）.

[5] Fukuda, K., Takayasu, M. and Takayasu, H.: A cause of self-similarity in TCP traffic, International Journal of Communication Systems, Vol.18, No.6, pp.603-617 (2005).

[6] 上田浩，奈須野裕，岩谷幸雄，木下哲男：確率過程による LAN トラヒックのモデル化における一考察，情報処理学会論文誌，Vol.48，No.SIG 2 (TOM 16), pp.167-174（2007）.

[7] Li, M.: Change trend of averaged Hurst parameter of traffic under DDOS flood Attacks, Computers & Security, Vol.25, No.3, pp.213-220 (2006).

[8] Hurst, H.E.: A suggested statistical model of some time series which occur in nature, Nature, 180, 494 (1957).

[9] Mandelbrot, B.B. and Wallis, J.R.: Robustness of the Rescaled Range R/S in the Measurement of Noncyclic Long-Run Statistical Dependence, Water Res., Vol.5, No.5, pp.967-988 (1969).

Akita University

[10] Beran, J., Sherman, R., Taqqu, M.S. and Willinger, W.: Long-Range Dependence in Variable-Bit Rate Video Traffic, IEEE Transactions on Communications, Vol.43, No.2/3/4, pp.1566-1579 (1995).

[11] Taqqu, M.S., Teverovsky, V. and Willinger, W.: Estimators for long-range dependence an empirical study, Fractals, Vol.3, No.4, pp.785-798 (1995).

[12] Igarashi, R., Ono, S., Takahashi, A., Iwaya, Y. and Sakata, M.: Some Features of Network Traffic Depending on Protocols, Proceedings of ICMR 2005, pp.426-431 (2005).

[13] Dang, T. and Molnar, S.: On the Effects of Non-Stationarity in Long-Range Dependence Tests, Periodica Polytechnica, Ser. El., Vol.43, No.4, pp.227-250 (1999).

[14] 松葉育雄：非線形時系列解析，pp.83-91，朝倉書店，東京 （2000）.

[15] Takahashi, A., Igarasjo, R., Ueda, H., Iwaya, Y. and Kinoshita, T.: Network Anomaly Detection Based on R/S Pox Diagram, International Journal of the Society of Materials Engineering for Resources, Vol.17, No.2, pp.186-192 (2010).

[16] 三鑰達真，吉田和幸：長期的スキャンニングを対象としたスキャン攻撃検知システム，電子情報通信学会技術研究報告，Vol.107，No.449，pp.39-44（2008）.

第2章　R/S Pox Diagram

2.1　はじめに

本章では，まず従来のR/S解析法[1]によるR/S Pox Diagramの導出手順について考察を行い，時系列データの経時変化においてトラフィック事象変化に対する即応性に関して問題が生じることを指摘し，その改善策を検討した．さらに，定常時とは異なるR/S Pox Diagramの特徴的プロット形状の発生要因について考察を行い，シミュレーション時系列を用いてその要因について検証を行った[2].

2.2　R/S解析法によるハーストパラメータ推定法

R/S解析法によるハーストパラメータ推定法の概要を図2.1に示す．まず，観測対象とする接続ポートに疎通するパケットトラフィックを計測するために，ネットワークスイッチにミラーポートを設定する．そのミラーポートに計測用のコンピュータを接続し, libpcap形式[3]パケット収集ツールを用いてパケットトラフィック時系列$X_t, t = 1, 2, \cdots, N$を作成する．具体的には，パケット収集ツールでパケットの到着時間を観測し，測定単位時間Δt毎の到着パケットを計数して得ることになる．libpcapツールを用いることにより，観測対象ポートの全パケットから特定のパケットのみフィルタリングした時系列データを生成することも可能となる．

この時系列X_tに対してR/S統計量を算出するための区間長$n_m, m = 1, 2, \cdots, M$の任意長区間を定める．n_mは$5 \leq n_m \leq N$の範囲で変動する区間長で，時系列X_tの系列長がNのとき，R/S統計量が算出される区間数mは，任意長区間が重複しない場合，$m = N/n_m$となる．例えば，$N = 3000$のとき，

- $n_m = 5$では，$m = N/n_m = 3000/5 = 600$
- $n_m = N/2 = 1500$では，$m = N/n_m = 3000/1500 = 2$
- $n_m = N = 3000$では，$m = N/n_m = 3000/3000 = 1$

などとなる．ただし，以後は記号の単純化のために，区間長n_mをnと記す．

ここで，観測時系列X_tの系列長Nが小さい場合，ハーストパラメータ推定の統計的な信頼性を高めるためには，任意長区間を重複させてR/S統計量のサンプル

Akita University

数を増やす必要がある [4] が，本手法ではハーストパラメータ自体の検証は行わないため，任意長区間は重複させずに R/S 統計量を求めた．

R/S 統計量の算出手順について説明する．

系列長 N の時系列 X_t で互いに重複せず，区間長が n となるような m 個の任意長区間を考える．この m 個の各区間において，以下のような統計量を求める．

まず，$1 \leq k \leq n$ なる k に対し，式 (2.1) より当該区間の区間平均 $\bar{X_n}$ を求める．

$$\bar{X_n} = \sum_{k=1}^{n} X_k / n \tag{2.1}$$

同一区間内での累積和と線形な傾向 $k\bar{X_n}$ との差を表す値 W_k を式 (2.2) より求める．

$$W_k = \sum_{j=1}^{k} X_j - k\bar{X_n} \tag{2.2}$$

さらに，式 (2.3) より，この W_k の最大値と最小値の差から累積範囲 R_n を求める．

$$\begin{aligned} R_n &= \max\{0, W_1, \cdots, W_k, \cdots, W_n\} \\ &\quad - \min\{0, W_1, \cdots, W_k, \cdots, W_n\} \end{aligned} \tag{2.3}$$

この累積範囲 R_n と式 (2.4) から求められる当該区間の標準偏差 S_n との比を用いて，R/S 統計量を導出する．

$$S_n = \sqrt{\sum_{k=1}^{n} X_k^2 / n - \bar{X_n}^2} \tag{2.4}$$

ここから，任意長区間 n を Δn ずつ増加させながらそれぞれの R/S 統計量を導出し，式 (2.5) に基づいてハーストパラメータ H を推定する．

$$\log(R_n/S_n) = H \log(n) + \log c \tag{2.5}$$

具体的には $\log(R_n/S_n)$ を被説明変数，$\log(n)$ を説明変数，$\log c$ を定数とした回帰モデルと想定し，最小 2 乗法によりハーストパラメータ H を推定する．これをグラフ的に表すと，横軸 $\log(n)$，縦軸 $\log(R_n/S_n)$ のグラフにプロットされた点群を表す回帰直線の傾きを求めることになる．このとき導出されるグラフが R/S Pox Diagram で，赤の破線が式 (2.5) で示される回帰直線を表す．

Akita University

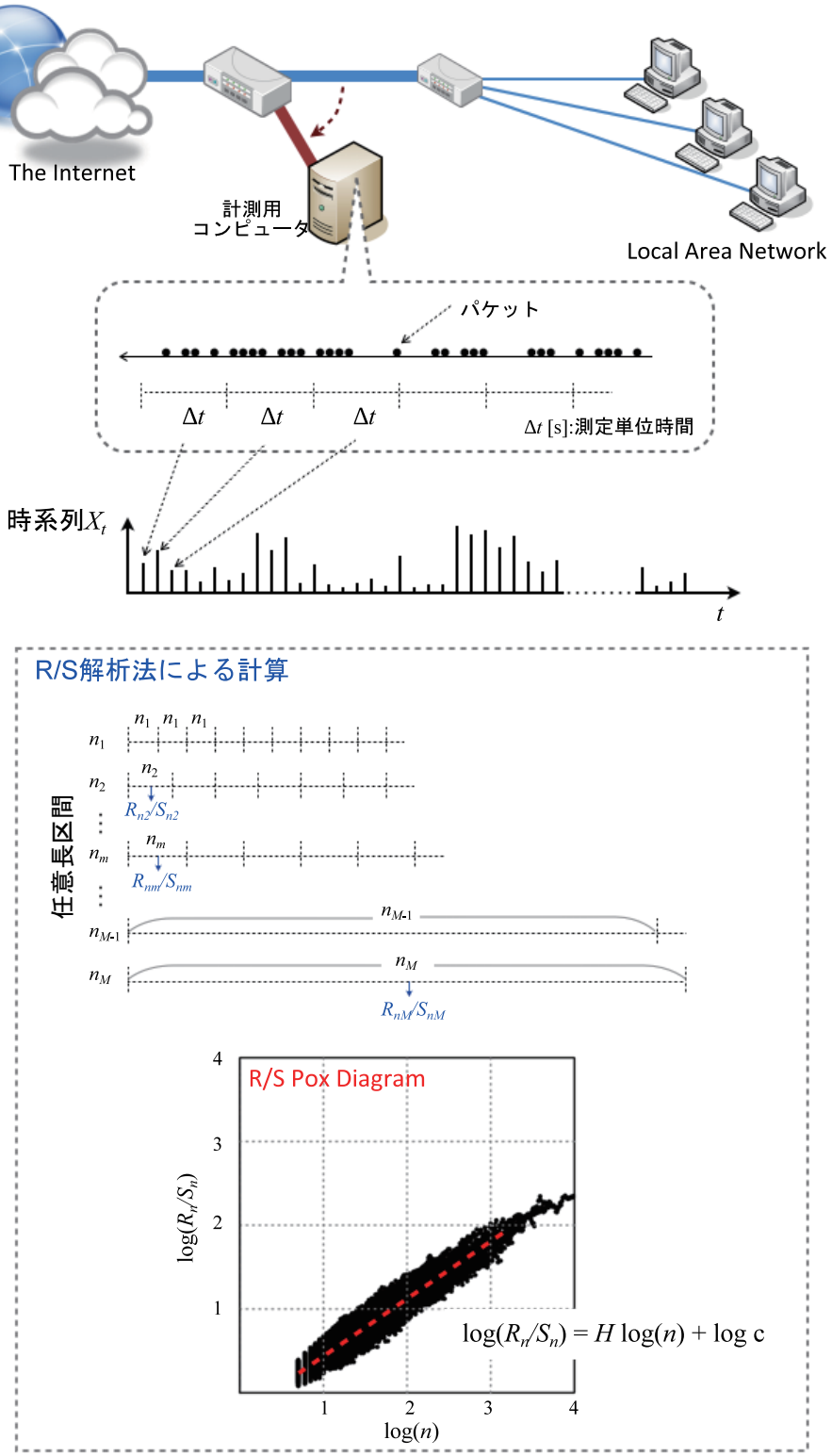


図 2.1 R/S 解析法によるハーストパラメータ H の導出

Akita University

2.3 R/S解析法に対する改良

2.3.1 R/S解析法の未計算区間

R/S解析法におけるトラフィック事象変化に対する即応性に関する問題点を指摘する．図2.2で示すように，R/S解析法では系列長Nの時系列X_tに対して区間長nの任意長区間を定め，その区間内のデータを用いてR/S統計量を算出する．区間長nは増分Δnずつ徐々に大きくして解析が行われるが，前項で言及したように任意長区間は重複させずに計算を行うため，系列長Nが区間長nで割り切れる場合，時系列X_t全区間データを用いてR/S統計量の計算が行われるが，割り切れない場合は時系列X_tの時間軸上後半の部分に計算されない区間が存在してしまう．この未計算区間は区間長nが大きくなるに従い，大きくなることが予想される．

通常，解析に用いる時系列データを作成するとき，観測時系列の時間軸は時間経過に伴い時刻tが大きくなることから，時刻tが小さい方向が「過去」のデータを表し，大きくなるにつれてより「直近」のデータを表すことになる．つまり，前述のR/S解析の未計算区間には，より「直近」のデータが存在することになり，トラフィック事象変化が解析に反映されにくいと考えられる．

上述の考察を検証するため，図2.2に示されている非定常的トラフィックが発生初期段階で観測されたと想定したシミュレーション時系列を用いてR/S Pox Diagramを導出した．そのときのR/S Pox Diagramを図2.3に示す．シミュレーション時系列X_tは，定常時の時系列として作成した系列長$N = 10000$のFGN(Fractional Gaussian Noize)に対し，区間[9901-10000]に振幅$P = 10$のレベルシフト系列を重畳させたものである．FGNは，R言語[5]のFGNパッケージ[6]にあるSimulateFGN関数を用いて，分散1.0，平均3.0，ハーストパラメータ$H = 0.5$と設定した．R/S Pox Diagramの灰色のプロット点は全任意長区間より算出されたR/S統計量を表し，青色のプロット点はパケット量が急激に増加したトラフィック変化点$t = 9901$が含まれる任意長区間から算出されたR/S統計量を表している．

結果より，トラフィック変化点が含まれる区間のプロット点は，全ての区間長nでプロットされるのではなく，推測どおり未計算区間の存在により一部の区間長nでは算出されずにまだらにプロットされることが確認できた．これは，DoS攻撃といった異常トラフィック等の観測対象によっては即応性が重視されることもあり，「直近」に発生したデータが解析に反映されにくいという現象は問題となる場合がある．

Akita University

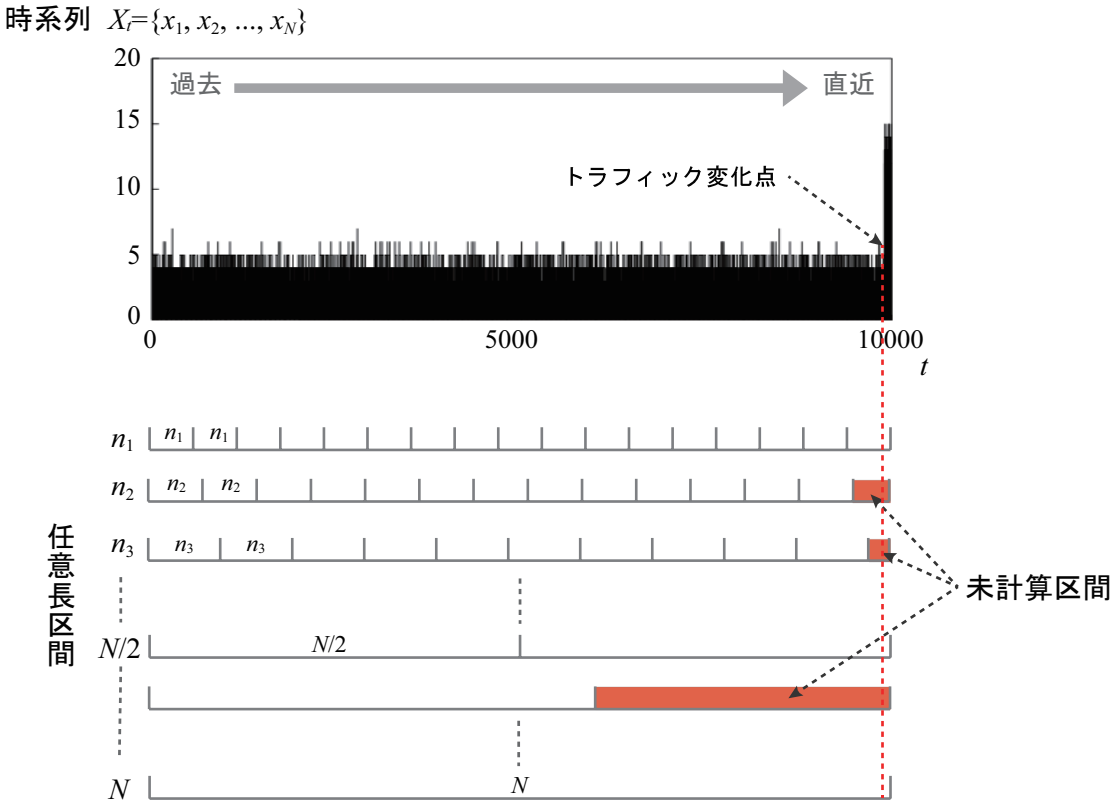


図 2.2 時系列 X_t と任意長区間の関係

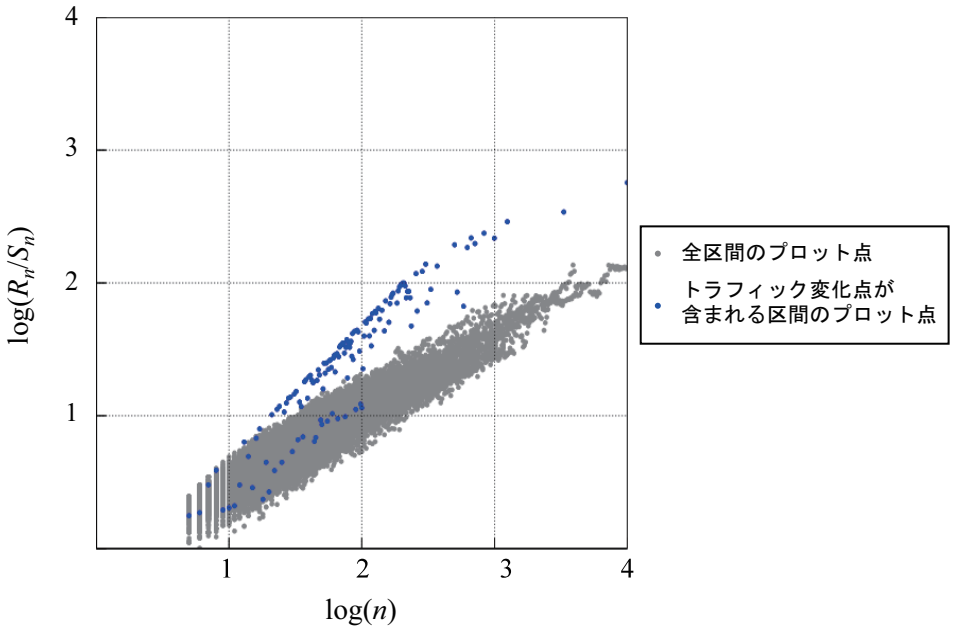


図 2.3 図 2.2 の時系列に対する R/S Pox Diagram

Akita University

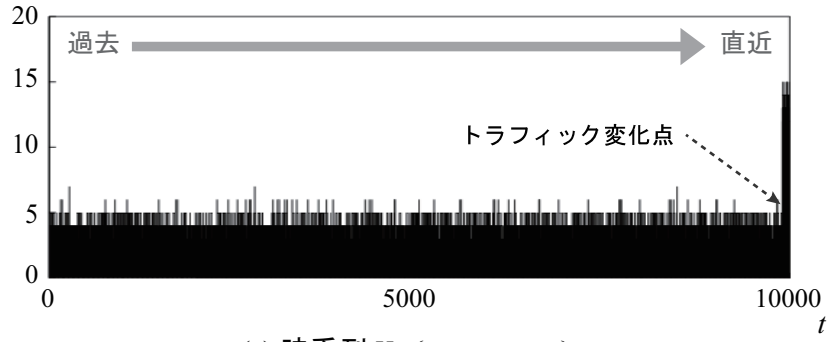
2.3.2 反転時系列による改良

この問題点を解決するためには，任意長区間を重複させてR/S統計量を算出することにより未計算区間をなくす方法が考えられる．しかし，この方法ではR/S統計量を算出する区間数が増えるため，計算コスト増加により即応性が低下する懸念が生じる．

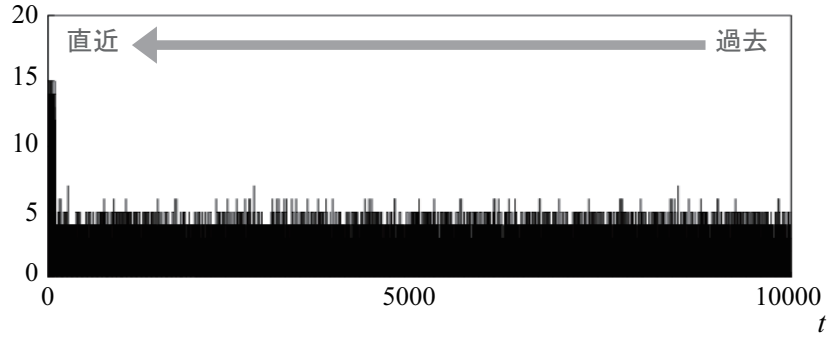
そこで，本論文では計算コストを増加させることなく解析を行える方法として，反転時系列による改良を提案する．従来の時系列 X_t を図2.4(a)，提案した反転時系列 X'_t を図2.4(b) に示す．これは，式(2.7) に示すように，時系列 X_t を単純に反転させたもので，より「直近」に近いデータを X'_t の前半に，「過去」のデータを後半にした時系列 X'_t として生成する．この改良法によって，トラフィック変化点が反転時系列 X'_t の前半 $1 \leq t \leq N/2$ の範囲に存在しているならば，全ての任意長区間 n において解析に反映されることになる．

$$X_t = \{x_1, x_2, \cdots, x_N\} \tag{2.6}$$

$$X'_t = \{x_N, \cdots, x_2, x_1\} \tag{2.7}$$



(a) 時系列 $X_t=\{x_1, x_2, ..., x_N\}$



(b) 反転時系列 $X'_t=\{x_N, x_{N-1},..., x_2, x_1\}$

図 2.4 時系列 X_t と反転時系列 X'_t

Akita University

図2.4(b) に示す反転時系列 X'_t を用いて導出したR/S Pox Diagram を図2.5 に示す．反転時系列 X'_t では全ての区間長 n に対してトラフィック変化点が計算に反映され，プロット点群が明確になることが観測できた．これより，反転時系列を用いることで「直近」で観測されるトラフィック変化に対して即応性が高まる効果が期待できる．

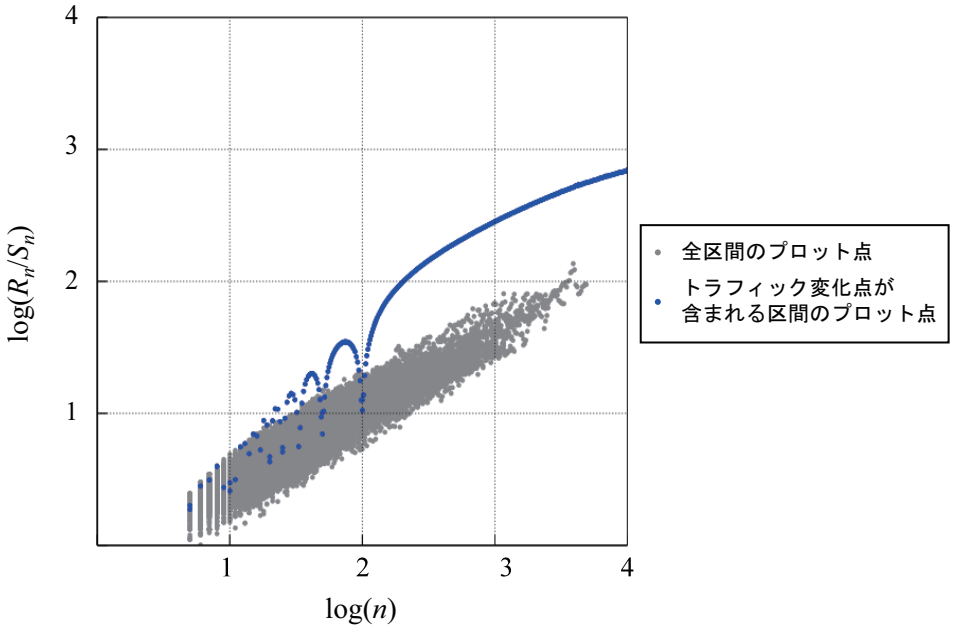


図 2.5 反転時系列に対するR/S Pox Diagram

Akita University

2.4 非定常性に対するR/S Pox Diagramプロット形状

R/S Pox Diagramのプロット形状は，定常状態のトラフィック時系列に対しては大体一定の傾きを呈するが，TCP SYN Flood攻撃[7]などの異常トラフィック時系列に対しては特徴的プロット形状を呈する場合がある[8][9]．特に，異常混入直後にはR/S Pox Diagramの上部にプロット点群が現れたり，ある点を境に途中から折れ曲がり二方向の傾きを示すプロット点群が現れたりする．これらの特徴は，定常状態からの変化を示す特徴と考えられるため，定量化することで変化検知の指標として利用可能と推測できる．そこで，この特徴の発生要因についてR/S解析の計算から考察し，シミュレーションにより検証する．本解析法に関する検証は，解析に用いるパラメータが多すぎて数学的検証が困難と考えられる．そこで，本研究ではシミュレーション時系列を用いた実験的検証を行った．

2.4.1 上部プロット点群の発生要因

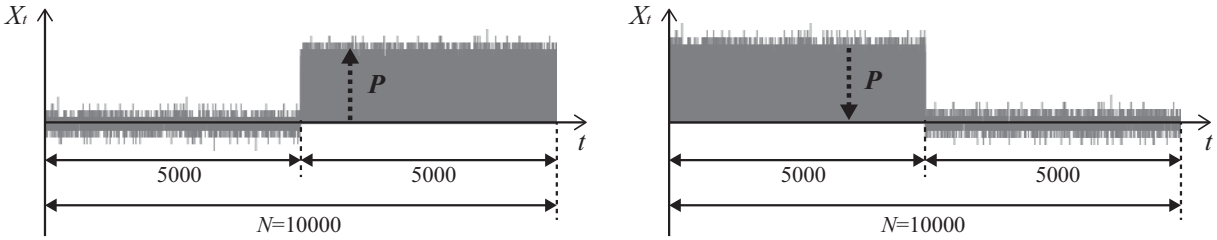
各任意長区間で現れる上部プロット点は，R/S統計量が大きくなる区間が発生することを意味する．R/S統計量は式(2.3)から求められる累積範囲R_nと式(2.4)から求められる標準偏差S_nの比より算出される．ここで，現時刻をtで表し，時刻tのとき算出された累積範囲を$R_n(t)$で表すと，R/S統計量が大きくなるということは，累積範囲R_nに式(2.8)で示す関係が成り立つ．

$$R_n(t) > R_n(t-1) \tag{2.8}$$

累積範囲R_nは当該区間の累積和と線形的な傾向との差を表す値W_k（式(2.2)）の最大値と最小値の差から算出されるので，W_kの値が大きくなる，つまり線形的な傾向より大きく外れる累積和になったとき，R/S統計量が大きくなることになる．トラフィック時系列ではパケット数が急激に増加，または減少したとき，この現象が現れると推測される．

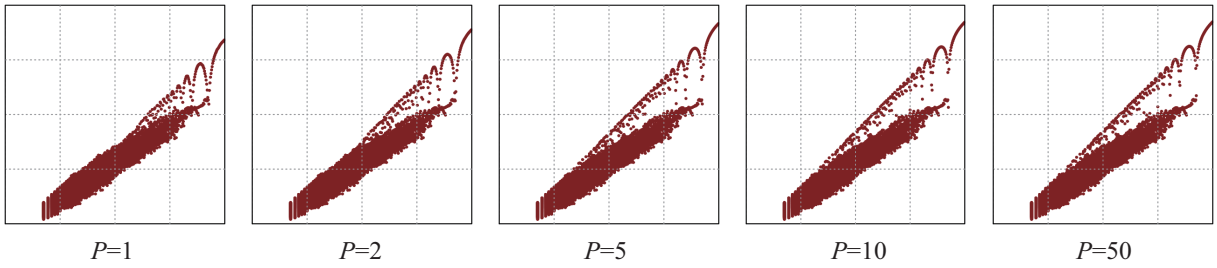
この推測を検証するために，突発的トラフィック量の増減を表す非定常性（レベルシフト時系列）を重畳させたシミュレーション時系列を用いてR/S Pox Diagramの導出を行った．検証に用いたトラフィック量増加を想定したシミュレーション時系列の概略図を図2.6(a)に，減少を想定したシミュレーション時系列を図2.6(b)に示す．系列長$N = 10000$のFGN（$H = 0.6$, 分散1.0, 平均0）に対して時刻$t = 5000$で振幅Pが変化する非定常時系列を重畳させた．シミュレーションは，増加・減少ともに振幅Pを1,2,5,10,50と変化させて行った．

Akita University

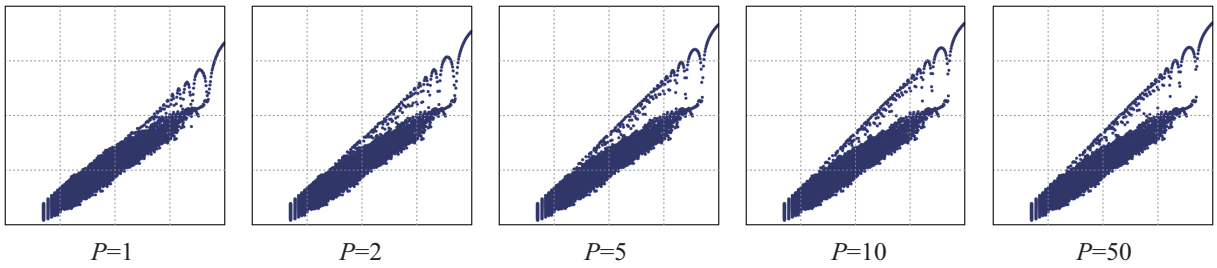


(a) レベルシフト（増加）　(b) レベルシフト（減少）

図 2.6 シミュレーション時系列（レベルシフト）



(a) レベルシフト（増加）



(b) レベルシフト（減少）

図 2.7 レベルシフトに対する R/S Pox Diagram

シミュレーション結果を図 2.7 に示す．結果より，振幅 P が 1 から 5 まで変化するとき，区間長 n が大きい範囲から上部に乖離するプロット点群が発生し，P が大きくなるにつれて，区間長 n が小さい範囲からもプロット点群が乖離していくことが観測できた．その後，P が 5 以上と変化させたときは，上部に乖離したプロット点群はほぼ同様のプロット形状を示すことも観測できた．図 2.7(a)，(b) とレベルシフトの増加・減少の方向性に関しては，ほとんど差異がないことも観測できた．ここから，上部プロット点群の発生要因は，レベルシフト変化を生じるような突発的トラフィック量変化にあることが実証できた．

Akita University

2.4.2 折れ曲がるプロット形状の発生要因

R/S Pox Diagramで折れ曲がるプロット形状の特徴として，変曲点のように見える任意長区間の区間長nでは，算出されるR/S統計量の範囲が小さくなり，ある値のR/S統計量に集約するようになる．さらに，このときの区間長nを境に前半・後半で異なる傾きを持つプロット形状を示す．R/S統計量がある値に集約するという特徴は，区間長nの任意長区間で算出される$m = N/n$個のR/S統計量がほぼ一定の値を示すということである．この場合，各区間内の時系列データが周期的にほぼ同様の変動パターンとなっていることが考えられる．

そこで，定常時系列に対して短時間にトラフィック量がパルス的に増加し，かつ周期的に発生するような周期的時系列を重畳させたシミュレーション時系列を用いてR/S Pox Diagramの導出を行った．検証に用いた周期的時系列の概略図を図2.8に示す．定常時系列は2.4.1と同様に系列長$N = 10000$のFGN（$H = 0.6$, 分散1.0, 平均0）を用いる．重畳させる非定常時系列は，系列長$N = 10000$, 振幅$P = 10$, パルス幅τ, 周期T（デューティ比$D = \tau/T$）のパラメータを設定した周期的時系列である．ここで，周期Tはパルス的パケットトラフィックの到着時間間隔を表すことになる．シミュレーションでは，周期Tを100，200，300と変化させ，パルス幅τはTに対してデューティ比$D = 50\%$となるように設定した．

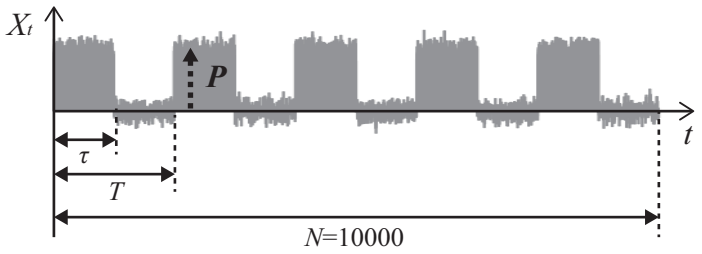


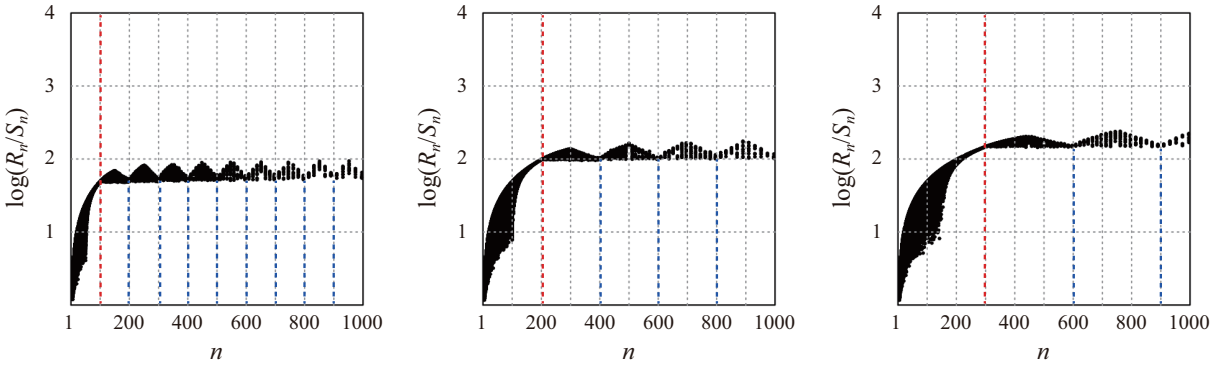
図2.8 シミュレーション時系列（周期的時系列）

Akita University

周期 $T = 100$ のときの R/S Pox Diagram を図 2.9(a)，周期 $T = 200$ を図 2.9(b)，周期 $T = 300$ を図 2.9(c) に示す．各グラフの区間長 n を表す横軸は実数軸で表し，$1 \leq n \leq 1000$ の範囲のみ拡大して表示している．赤い破線は設定した周期 T と同じ値の区間長 n を表し，青い破線は $n = lT, l = 2, 3, \cdots$ と周期 T を整数倍したときの区間長 n を表す．

結果より，どの周期 T においても周期 T と同じ値の区間長 n のとき，求められた R/S 統計量はほぼ同じ値に集約することが確認できた．さらに，$T = 100$ のときは，$n = 200, 300, \cdots$，$T = 200$ のときは，$n = 400, 600, \cdots$，$T = 300$ のときは，$n = 600, 900, \cdots$ のように，区間長 n の整数倍となる区間長 ln のとき，同様に R/S 統計量がある値に集約していることも観測された．これは，区間長 n の各時系列データの特徴がほぼ同じことから，区間長 n が整数倍となっても統計的特徴は変化しないことが要因と推測できる．つまり，$n \geq T$ の範囲では，R/S 統計量がある値に集約するのを繰り返すことで，プロット形状は水平方向に傾くことになる．よって，折れ曲がるプロット形状の発生要因は，周期的時系列によるものと実証できた．

ここから，折れ曲がるプロット形状の特徴を定量化することにより，トラフィック事象における周期的特徴を示す検知指標として利用できると推測される．



(a) 周期 T=100　(b) 周期 T=200　(c) 周期 T=300

図 2.9 周期的時系列に対する R/S Pox Diagram

Akita University

2.5 まとめ

本章では，R/S 解析法による R/S Pox Diagram のトラフィック変化に対する即応性に関する改善策の有用性について検討を加えた．さらに，非定常性に対する R/S Pox Diagram の特徴的プロット形状の発生要因に関してシミュレーションによる実験的検証を行った．これらにより，得られた成果を以下にまとめる．

(1) R/S 解析法に対して反転時系列 X_t' を用いることで，「直近」で発生したトラフィック変化が未計算区間に含まれることなく，全任意長区間で反映され，即応性が高まることを明らかにした．
(2) R/S Pox Diagram の上部に発生するプロット点群は，突発的にトラフィック量が増減するようなレベルシフト時系列に対して発生することを明らかにした．
(3) R/S Pox Diagram の折れ曲がるようなプロット形状は，周期的にパルス的トラフィックが到着するような周期的時系列に対して発生することを明らかにした．

参考文献

[1] Leland, W.E., Taqqu, M.S., Willinger, W. and Wilson, D.V.: On the Self -Similar Nature of Ethernet Traffic, Computer Communications Review, Vol.23, No.4, pp.183-193 (1993).

[2] 高橋秋典，五十嵐隆治，上田浩，岩谷幸雄，木下哲男：R/S Pox レッグライン特性，情報処理，Vol.54，No.6，pp.1761-1770 (2013).

[3] TCPDUMP&LiBPCAP, http://www.tcpdump.org/

[4] 松葉育雄：長期記憶課程の統計，共立出版 (2007).

[5] The R Project for Statistical Computing, http://www.r-project.org/

[6] FGN: Fractional Gaussian Noize and hyperbolic decay time series model fitting, http://cran.r-project.org/web/packages/FGN/

[7] CERT Advisory CA-1996-21: TCP SYN Flooding and IP Spoofing Attacks, http://www.cert.org/historical/advisories/CA-1996-21.cfm

[8] 高橋秋典, 五十嵐隆治, 上田浩, 岩谷幸雄, 木下哲男：R/S Pox Diagram に基づくトラフィック異常検知に関する研究, 電子情報通信学会技術研究報告, Vol.108, No.203, pp.45-50 (2013).

[9] Takahashi, A., Igarasjo, R., Ueda, H., Iwaya, Y. and Kinoshita, T.: Network Anomaly Detection Based on R/S Pox Diagram, International Journal of the Society of Materials Engineering for Resources, Vol.17, No.2, pp.186-192 (2010).

第3章　R/S Pox レッグライン特性

3.1　はじめに

本章では，R/S Pox Diagram に現れる特徴的プロット形状を定量化し，非定常性の存在を示す指標となる新たな特性を提案する．この特徴的プロット形状には，人体脚部の太腿，膝，脛と非常によく似ている形状を呈する場合がある．そこで，本論文では，この類似性から R/S Pox レッグライン (Leg-Line) 特性と呼ぶことにした．また，この特性を定量化した特徴量について定義を行い，この特徴量を用いた周期的時系列における周期 T を推定する手法について提案する．さらに，提案手法の有効性を示すため，シミュレーション時系列に対する特徴量の性能評価，および周期推定実験を行った [1].

3.2　特徴量の定義

提案する R/S Pox レッグライン特性の特徴量を図 3.1 に示す．折れ曲がるプロット形状の特徴は，脚部の太腿 (Thigh) と脛 (Shin) の部分に見立て，ハーストパラメータ H 導出法と同様にそれぞれの部分におけるプロット点群の傾きで表す．太腿に見立てた区間長 n が小さい範囲を導出範囲 RT(Range of Thigh)，区間長 n が大きい範囲を導出範囲 RS(Range of Shin) として，それぞれの範囲を特徴付けるプロット点群の傾き（Slope 値）を ST(Slope of Thigh)，SS(Slope of Shin) と表記する．また，周期的時系列の周期 T と推定されるプロット点群が折れ曲がるときの区間長 n を，脚部の膝 (Knee) と見立て，KP(Knee Point) と表記する．

導出範囲 RT および RS におけるプロット点群は，図 2.5 に示されるようにトラフィック変化点が含まれない定常時には一定方向の傾きを示すが，ある程度の幅を有したプロット形状を示す．トラフィック変化が生じて，上部にプロット点群が発生したときも定常時のプロット形状は経過時間に対してある程度残存すると推測される．このように，様々な傾斜の形状を呈していることから，各範囲において一方向の傾きではなく，プロット点の各区間長 n における上限点群 $\max(R_n/S_n)$，平均点群 $\mathrm{avg}(R_n/S_n)$，下限点群 $\min(R_n/S_n)$ から求まる三方向の傾きを算出して形状の特徴を示すことにした．このとき，導出範囲 RT におけるそれぞれの傾きを $\{ST_{\mathrm{Sup}}$,

Akita University

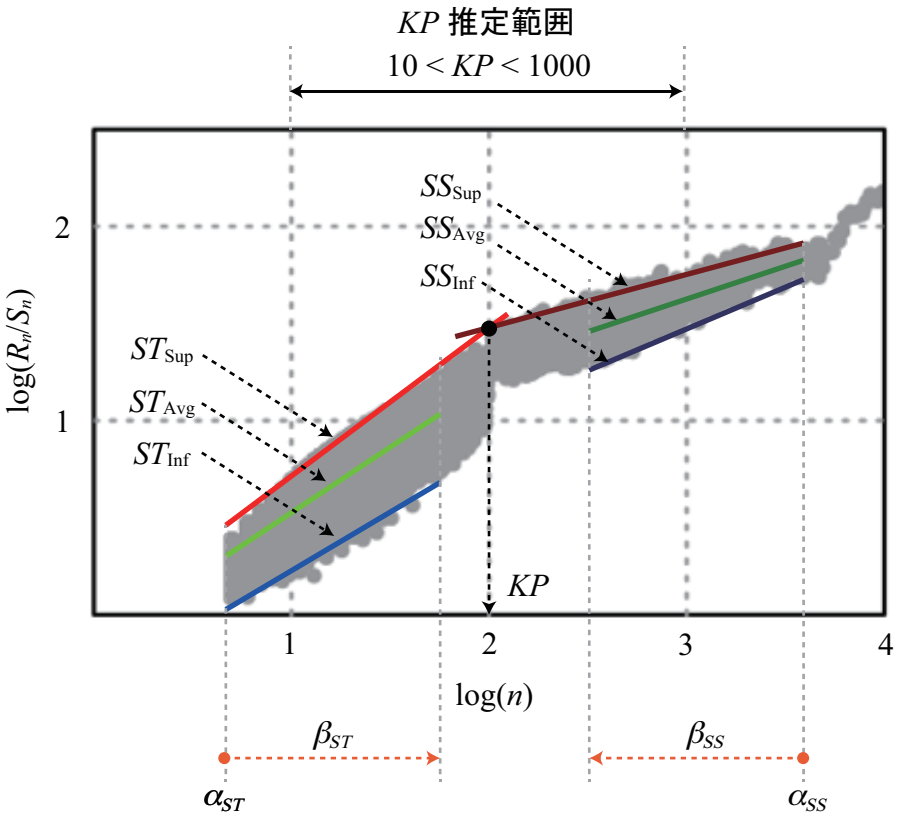


図 3.1 R/S Pox レッグライン特性

ST_{Avg}，$ST_{\mathrm{Inf}}\}$，導出範囲 RS におけるそれぞれの傾きを $\{SS_{\mathrm{Sup}}$，SS_{Avg}，$SS_{\mathrm{Inf}}\}$ と表記し，それぞれの値はハーストパラメータ H を求めた式 (2.5) と同様にそれぞれのプロット点群に対して最小 2 乗法を用いて推定する．導出範囲 RT，RS は，図 3.1 において区間長 n に対してそれぞれ起点 α_{ST}，および α_{SS} を設定し，$\log(n)$ 上の値で RT および RS の範囲長 β_{ST}，β_{SS} を設定する．従って RT は α_{ST} を起点とし範囲長 β_{ST}の範囲，RS は α_{SS} を起点とし，図示の向きの範囲長 β_{SS} の範囲とする．

本論文における導出範囲 RT，RS の起点 α_{ST}，α_{SS} および範囲長 β_{ST}，β_{SS} の設定について具体的に説明する．

まず，RT の起点 α_{ST} は，R/S 統計量を算出するためのサンプル数確保という観点から区間長 $n = 5$，つまり $\alpha_{\mathrm{ST}} = \log(5)$ とし，RS の起点 α_{SS} は，区間長 n が大きくなると求められる R/S 統計量の数が少なくなるため，上限点群，平均点群，下限点群の傾きを求めるという観点から最低三点が求められる区間長 $n = N/3$，つまり $\alpha_{\mathrm{SS}} = \log(N/3)$ と設定した．

次に，導出範囲 RT，RS の範囲長 β_{ST}，β_{SS} は，サンプル数とのトレードオフになる．すなわち，図 3.1 に示した各量の統計的新再生を高める場合，RT ないしは RS の範囲長を大きくする必要があるが，この範囲長を大きくし過ぎるとそれぞれの

Akita University

導出範囲をカバーしてしまい，後述する周期推定に影響すると懸念される．また，図2.8に示すように，周期Tの変動によってKPも変動するため，極端な短周期および長周期を精度高く推定するためには，それぞれの場合に対する範囲長β_{ST}，β_{SS}を調整する必要がある．しかし，混在する周期的時系列の周期Tは未知であることから，本論文では観測対象から推定可能となる周期Tの範囲を限定し，その範囲において周期推定を行うことを目標とし，RT，RTの範囲長を検討した．この方法では，周期推定範囲は限られるが，観測時系列X_tを生成するときの測定単位時間Δtを調整すれば，実時間上で異なる周期について推定が可能となる．つまり，並列処理により広範囲または限定した周期的時系列の観測システムへの適用も検討できる．本論文では，実時間解析における計算コストを考慮したときの時系列X_tの系列長Nを10000以下として検討することから，有意な推定範囲として$10 \leq T \leq 1000$を目標とした．この目標とした範囲内で周期推定を行える範囲長β_{ST}，β_{SS}として，本論文では経験的にそれぞれ0.8と設定した．

Akita University

3.3 Knee Pointによる周期推定アルゴリズム

図3.2に示すKnee Pointによる周期推定アルゴリズムについて説明する．周期推定法の基本的なアイデアは，導出範囲RT, RSから求められた各傾きST_{mode}, SS_{mode}を表す回帰直線の交点となる区間長nを推定周期KPとして求めるものである．ここで，modeとはそれぞれの導出範囲における上限点群Sup，平均点群Avg，下限点群Infを表す．

まず，導出範囲RT，RSから交点を求めるための回帰直線ST_{mode}およびSS_{mode}をそれぞれ一つずつ選択する．この回帰直線を用いて交点を算出するが，それぞれの回帰直線の傾きが平行でない場合は常に交点が存在することになる．つまり，周期的時系列が重畳されていなくても，推定周期となるKPが導出されてしまい，R/S Pox Diagramの折れ曲がるプロット形状と異なる点が周期として推定されてしまうことになる．この不具合を回避するために，周期的時系列が重畳されるとSS_{mode}の値が小さくなることから，SS_{mode}に対して周期性判定となるしきい値γを定め，SS_{mode}の値がしきい値γより小さくなったときのみ，交点を計算するという制約条件を設けた．

交点の計算について説明する．回帰直線ST_{mode}を式(3.1)，回帰直線SS_{mode}を式(3.2)と表す．ここで，c_{ST}およびc_{SS}はそれぞれの回帰直線の切片を表す．

$$\log(R_n/S_n) = ST_{\mathrm{mode}} \log(n) + c_{\mathrm{ST}} \tag{3.1}$$

$$\log(R_n/S_n) = SS_{\mathrm{mode}} \log(n) + c_{\mathrm{SS}} \tag{3.2}$$

これより，2直線の交点のx座標はそれぞれ式(3.3)より求められる．

$$x = \log(n) = \frac{c_{\mathrm{ST}} - c_{\mathrm{SS}}}{SS_{\mathrm{mode}} - ST_{\mathrm{mode}}} \tag{3.3}$$

この$x = \log(n)$，つまり区間長nが周期Tを表現することから，この値を推定周期KPとして求める．ただし，KPの値が推定目標とした$10 \leq T \leq 1000$の範囲外となったとき，算出できなかったものとする．

STおよびSSのmodeはそれぞれ3つあるので，ここから求められるKPの数は9個となる．それぞれの傾きより算出されるKPを表3.1に示す．

Akita University

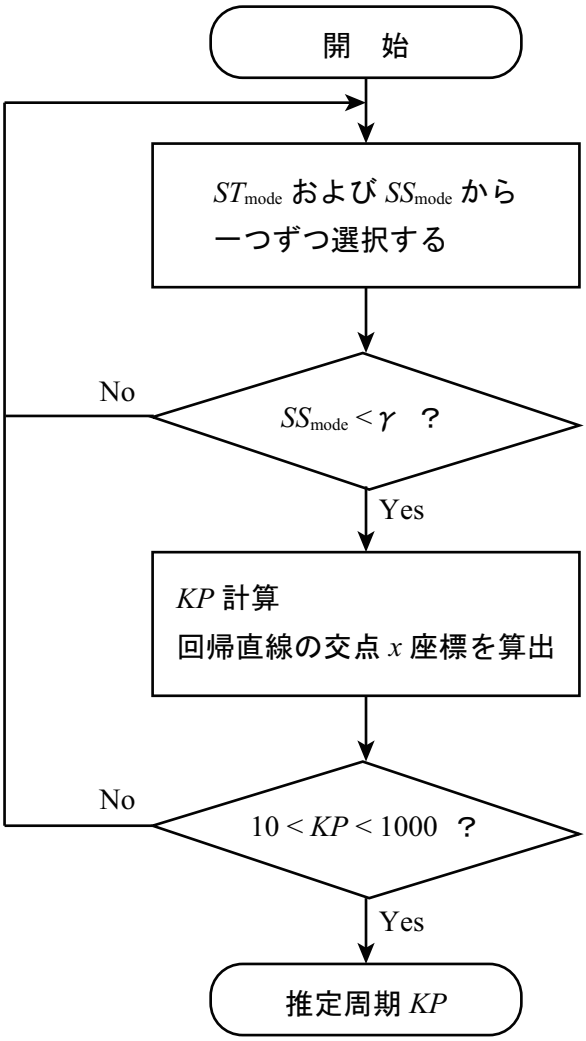


図 3.2 KP による周期推定アルゴリズム

表 3.1 導出される KP

	SS_{Sup}	SS_{Avg}	SS_{Inf}
ST_{Sup}	KP_{SS}	KP_{SA}	KP_{SI}
ST_{Avg}	KP_{AS}	KP_{AA}	KP_{AI}
ST_{Inf}	KP_{IS}	KP_{IA}	KP_{II}

Akita University

3.4 シミュレーション時系列による性能評価

提案するR/S Poxレッグライン特性の特徴量が，非定常性に対して顕現するR/S Pox Diagramの特徴的プロット形状を表現しているかを評価するため，シミュレーション時系列を用いて特徴量を算出して検証を行った．具体的には，2.4で用いたレベルシフト時系列，および周期的時系列において各時系列のパラメータ（振幅P，周期T，デューティ比D）を変化させたときの特徴量（Slope値）の変化を，以下に示す各特性に基づいてシミュレーションにより検証した．

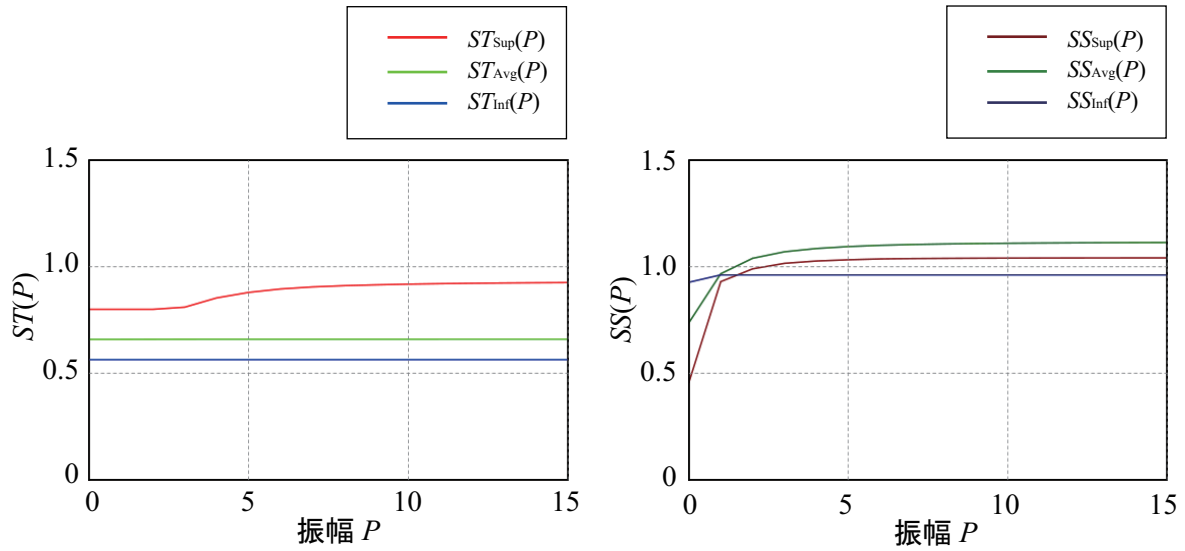
(1) 振幅特性：レベルシフト時系列のトラフィック量変化に対する特徴量
(2) 周期特性：周期的時系列のパケット到着時間間隔変化に対する特徴量
(3) 周期ー振幅特性：周期的時系列のトラフィック量および周期に対する特徴量
(4) デューティ比ー周期特性：周期的に到着するパケット量の変化に対する特徴量
(5) 時間特性：非定常時系列混入進捗に対する特徴量の時間変化

3.4.1 振幅特性

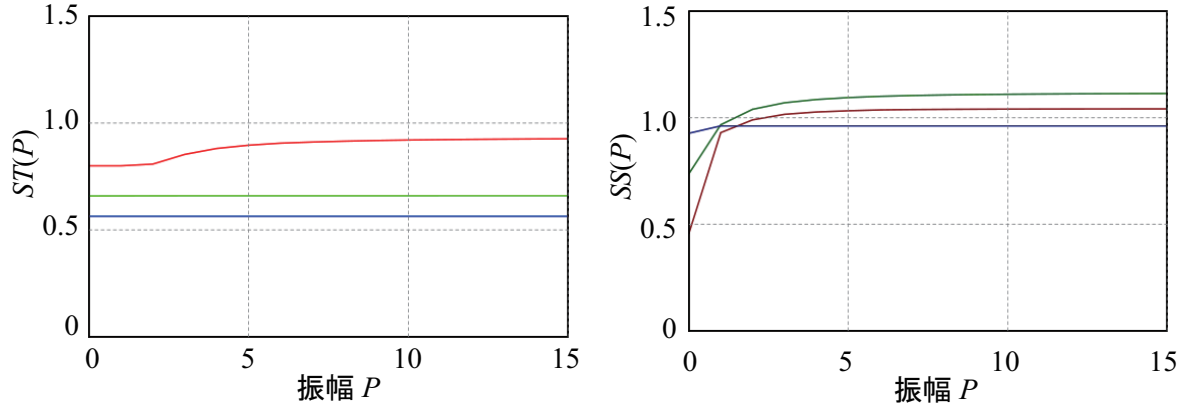
振幅特性のシミュレーション概要について説明する．このシミュレーションはパケットトラフィックにおける突発的トラフィック量の増減を想定したもので，レベルシフト量（振幅）に対する特徴量の変化を検証する．シミュレーション時系列は，図2.6(a),(b)で示される増加・減少を表すレベルシフト時系列と同様に，系列長$N = 10000$，ハーストパラメータ$H = 0.6$，分散1.0，平均0のFGN時系列に対して時刻$t = 5000$で振幅Pを変化させたものを用いる．この振幅Pを$0 \leq P \leq 15$と変化させたときのシミュレーション時系列に対するR/S Poxレッグライン特性のSlope値を算出した．

レベルシフト時系列（増加）に対する結果を図3.3(a)に示す．結果より，振幅Pの増加に伴い，上限点群から求められる$ST_{\mathrm{Sup}}(P)$および$SS_{\mathrm{Sup}}(P)$の値が高くなる傾向が示された．ただし，$P > 5$になるとそれぞれの値の変動はなくなり，ある値に収束することがわかった．$ST_{\mathrm{Avg}}(P)$，$ST_{\mathrm{Inf}}(P)$は振幅Pの影響は受けずに変化を示さなかった．SSにおいては，$SS_{\mathrm{Avg}}(P)$，$SS_{\mathrm{Inf}}(P)$も値が高くなる傾向を示したが，これは，STに比べR/S統計量のサンプル数が少ないための効果と考えられる．この結果より，レベルシフト（増加）の発生による上部プロット点群の発生はST_{Sup}およびSS_{Sup}で検知可能と推測される．

Akita University



(a) レベルシフト時系列（増加）



(b) レベルシフト時系列（減少）

図 3.3 レベルシフト時系列による振幅特性

図 3.3(b) に示すレベルシフト時系列（減少）に対する結果より，各 Slope 値の変動傾向は増加のレベルシフト時系列とほぼ同様の傾向を示したことから，減少のレベルシフトも検知可能と推測される．ただし，増加，減少の方向性を認識できないことから，その判別に対する方法も検討が必要と考えられる．

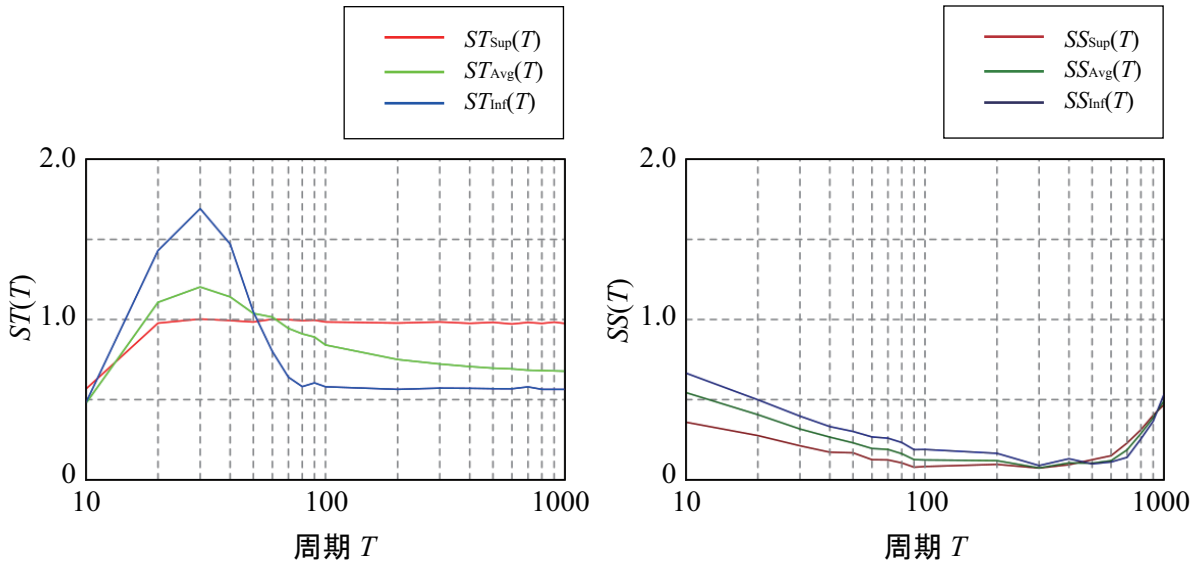
Akita University

3.4.2 周期特性

パケットトラフィックにおける周期性は，ある一定の時間間隔をおいて到着するような場合に発生するもので，Pulsing DoS 攻撃 [2][3] のように非常に短い時間で比較的少数のパケットを断続的に送信する場合や，攻撃検知をしにくくするために長期間に調査パケットを送信するような長期的ポートスキャン [4] が発生している場合などが想定される．つまり，周期性を有する通信には，ある意図を持って送信される場合が多いことから，周期性の存在を検知することは有効と考えられる．

周期特性のシミュレーション概要について説明する．シミュレーション時系列は，図 2.8 で示される周期的時系列を用いて，振幅 $P = 10$，デューティ比 $D = 50\%$ としたとき，周期 T を $10 \leq T \leq 1000$ と変化させたときの Slope 値を算出した．

周期 T に対する ST の結果を図 3.4(a)，SS の結果を図 3.4(b) に示す．ST において ST_{Sup} は $20 < T \leq 1000$ のとき，安定して約 1.0 の値を示している．ST_{Avg} および ST_{Inf} は $T = 30$ のとき最大値をとり，その後 T が増加するにつれて，値が低下している．特に ST_{Inf} は，$T = 100$ 以降，定常状態とほぼ同じ値に戻っている．以上より，ST_{Sup} は広い範囲の周期 T に対しても安定した周期的時系列の検知ができると推測できる．



(a) Slope of Thigh　　(b) Slope of Shin

図 3.4 周期 T に対する Slope 値

Akita University

SSでは，どの傾きも$20 \leq T \leq 800$において0.3以下の値となり，水平方向に傾く特徴を捉えることができた．ただし，Tが1000近辺になると傾きが元に戻るようになる．これは，RSの導出範囲設定に起因するもので，β_{SS}の調整および系列長Nの設定により，SSの変動傾向を示す周期Tの範囲を変更できると推測される．

3.4.3 周期ー振幅特性

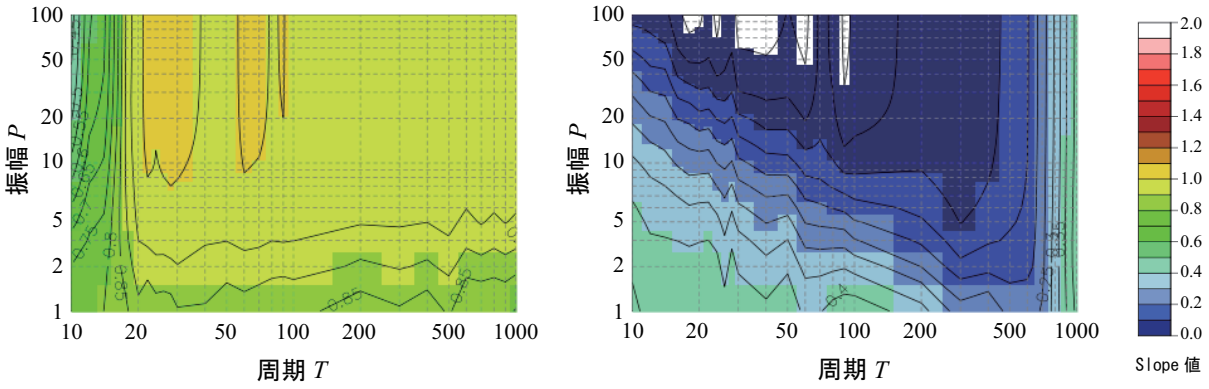
周期ー振幅特性のシミュレーション概要について説明する．周期特性と同時に振幅Pの変化をシミュレートすることにより，周期的トラフィックのパケット量に対する性能を確認することができる．シミュレーション時系列は，周期特性と同様に図2.8で示される周期的時系列を用いて，デューティ比$D = 50\%$としたとき，周期Tを$10 \leq T \leq 1000$，振幅Pを$1 \leq P \leq 100$と変化させたときのSlope値を算出した．

シミュレーション結果を図3.5に示す．それぞれ図3.5(a)は$ST_{\mathrm{Sup}}(T, P)$，図3.5(b)は$SS_{\mathrm{Sup}}(T, P)$，図3.5(c)は$ST_{\mathrm{Avg}}(T, P)$，図3.5(d)は$SS_{\mathrm{Avg}}(T, P)$，図3.5(e)は$ST_{\mathrm{Inf}}(T, P)$，図3.5(f)は$SS_{\mathrm{Inf}}(T, P)$を示している．各グラフは，横軸に周期T，縦軸に振幅Pを表した両対数グラフで，Slope値は数値に対してカラーを対応させて表示している．カラー表示は，定常時のSlope値に対してあまり変化が無い場合は緑色を示し，Slope値が高くなると黄から赤へ，低くなると青に変化するように表現されている．ただし，図3.5(b)および図3.5(c)にある白の領域は1.9〜2.0の高い値を示しているのではなく，負の値となって描画色が無い状態が示されている．

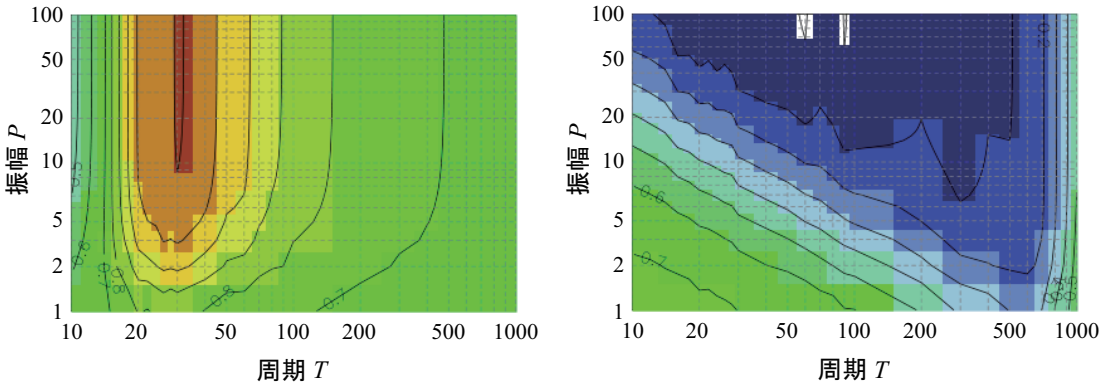
結果より，振幅Pが5以上の場合，周期特性で観測されたように，各STとも周期Tが大きくなるにつれて値が上昇し，ST_{Sup}は周期Tが20以上となると安定して約1.0の値となり，ST_{Avg}，ST_{Inf}は周期Tが100以上となると定常時とほぼ同じ値となった．これは，レベルシフト時系列の振幅特性と同様の性質を示す結果となった．つまり，定常状態に対して周期的時系列の振幅Pがある程度の大きさを有していれば，振幅Pの大きさはSlope値の変動傾向に対して影響を及ぼさないというロバスト性を有していると推測できる．このとき，振幅Pが変化してもST_{Sup}は約1.0の値を保っていることから，検知指標としての有効性が推測される．

SSにおいては，SS_{Sup}はSS_{Avg}，SS_{Inf}に比べ振幅P，周期Tの広範囲で値が小さくなることから，それぞれの変化に対してロバスト性を有していると推測される．しかし，SS_{Avg}，SS_{Inf}は小さい振幅Pおよび短い周期Tに対して感度が鈍いことが観測できた．よって，周期性の判定指標として利用する場合，判定が困難となることも懸念される．

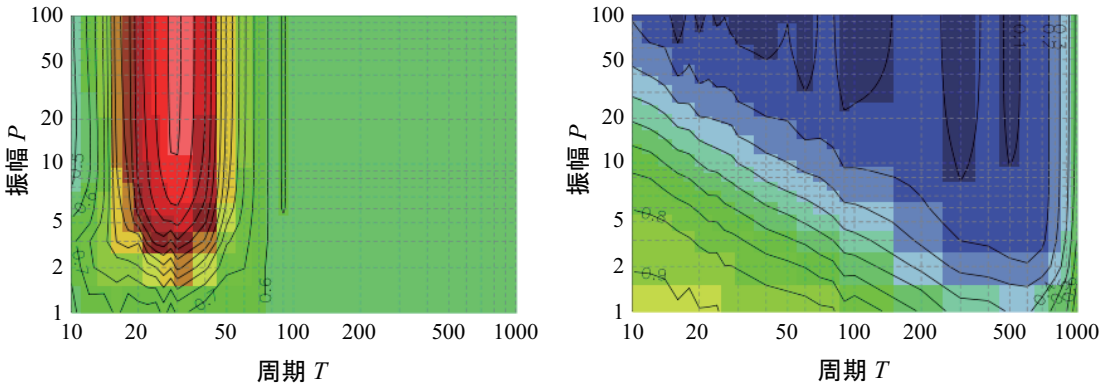
Akita University



(a) $ST_{Sup}(T, P)$ (b) $SS_{Sup}(T, P)$



(c) $ST_{Avg}(T, P)$ (d) $SS_{Avg}(T, P)$



(e) $ST_{Inf}(T, P)$ (f) $SS_{Inf}(T, P)$

図 3.5 周期的時系列による周期-振幅特性

Akita University

3.4.4 デューティ比ー周期特性

デューティ比ー周期特性のシミュレーション概要について説明する．デューティ比は周期Tに対するパルス幅τの影響，つまり周期ー振幅特性と同様に周期的トラフィックのパケット量に対する性能を確認することができる．デューティ比が小さいとパケット量が少なく，大きいとパケット量が多い周期的トラフィックとなる．シミュレーション時系列は，図2.8で示される周期的時系列を用いて，振幅$P = 10$としたとき，デュティ比Dを$0 < D < 1.0(100\%)$，周期Tを$10 \leq T \leq 1000$と変化させたときのSlope値を算出した．

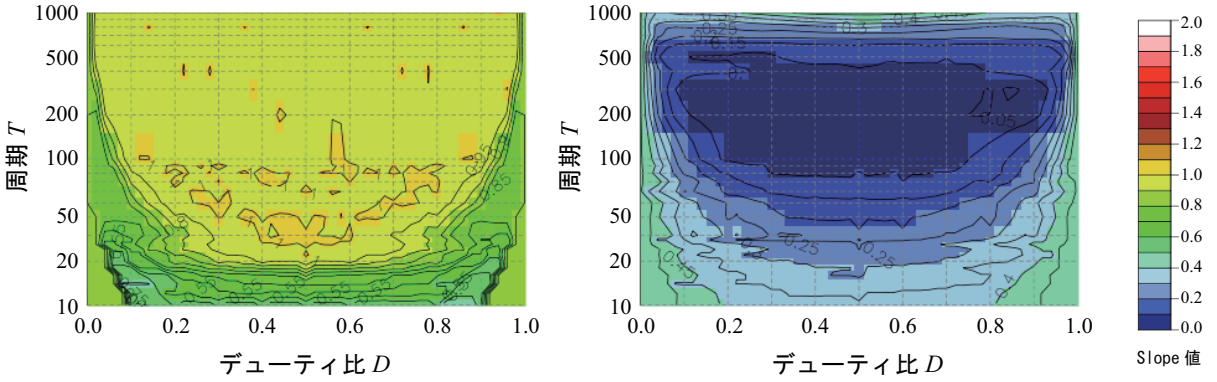
シミュレーション結果を図3.6に示す．それぞれ図3.6(a)は$ST_{\mathrm{Sup}}(D,T)$,図3.6(b)は$SS_{\mathrm{Sup}}(D,T)$, 図3.6(c)は$ST_{\mathrm{Avg}}(D,T)$, 図3.6(d)は$SS_{\mathrm{Avg}}(D,T)$, 図3.6(e)は$ST_{\mathrm{Inf}}(D,T)$, 図3.6(f)は$SS_{\mathrm{Inf}}(D,T)$を示している．各グラフのカラー表示は周期ー振幅特性と同様である．

結果より，STおよびSSのどちらの場合においても，デューティ比$D = 50\%$においてそれぞれほぼ最大値および最小値をとることが観測された．つまり，このときの値が周期的時系列の特徴を最もとらえた値と推測される．また，各値とも$D = 50\%$を中心にDが大きく，または小さくなるにつれて，定常時からの変化は小さくなることが観測された．しかし，その傾向は周期Tが大きくなるにつれて小さくなることも観測された．この傾向から，周期Tとデューティ比Dには特徴量に対して何らかの関係性を有していると推測される．

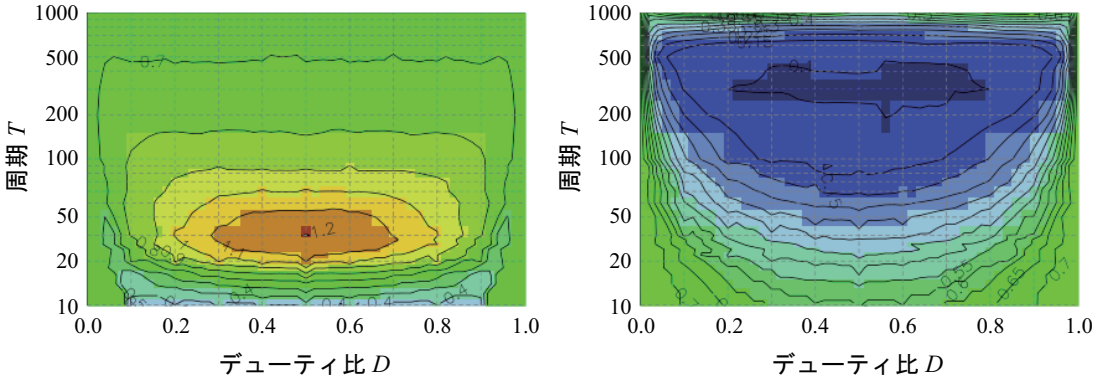
ここで，デューティ比Dと周期Tの関係性を調べるため，図3.6(b)より，各周期Tにおいてデューティ比Dを徐々に増やしたときに，Slope値SS_{Sup}が周期判定しきい値$\gamma = 0.3$以下となった時点でのパルス幅τを調査した．その結果を図3.7に示す．赤い破線は，求められたパルス幅τの平均値を表す．

結果より，周期Tが変化してもしきい値γ以下となるパルス幅τは，平均値6.82, 約7ポイントであることが観測された．つまり，本シミュレーションの場合，周期的時系列のパルス幅が7ポイント以上あれば，デューティ比に影響されること無く，周期性を判定できることが推測される．

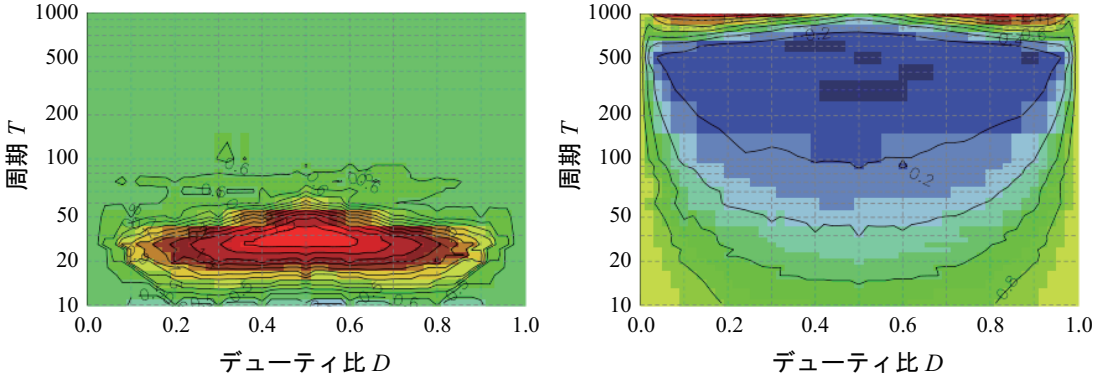
Akita University



(a) $ST_{Sup}(D, T)$ (b) $SS_{Sup}(D, T)$



(c) $ST_{Avg}(D, T)$ (d) $SS_{Avg}(D, T)$



(e) $ST_{Inf}(D, T)$ (f) $SS_{Inf}(D, T)$

図 3.6 周期的時系列によるデューティ比-周期特性

Akita University

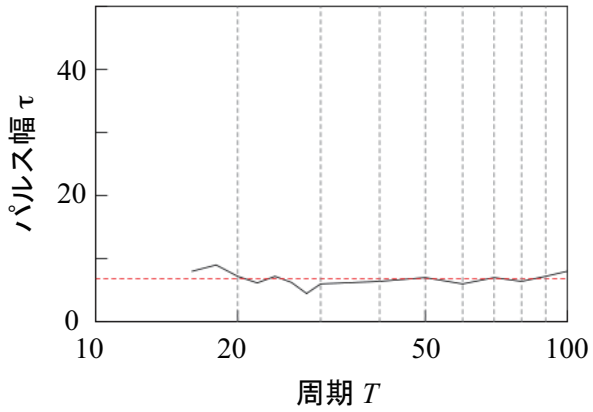


図 3.7 周期T対するSS_{Sup}が周期判定しきい値γ以下となるパルス幅τ

Akita University

3.4.5 時間特性

実時間上の時系列特性解析では，非定常性を有するトラフィック時系列が発生した場合，観測対象となる時系列 X_t に対して徐々に混入してくることになる．前述までの特性解析では，レベルシフト時系列は時系列 X_t の半分まで，また周期的時系列は時系列 X_t の全体に混入したと想定していたが，特徴量の経時変化を観測するため，非定常時系列の重畳系列長を変化させた場合の時間特性を検証した．時間特性シミュレーションの概要を図3.8に示す．観測時系列 X_t に対して徐々に重畳される非定常時系列は，重畳される非定常時系列の系列長 L_t で経時変化を表現している．本シミュレーションでは，時系列 X_t に対して重畳ステップサイズ ΔL_t を 10 ずつ増加させながら非定常時系列を重畳させた観測時系列を用いて特徴量を算出した．非定常時系列は，振幅 $P = 10$ のレベルシフト時系列，および振幅 $P = 10$，周期 $T = 100$，デューティ比 $D = 50\%$ の周期的時系列を用いて検証を行った．

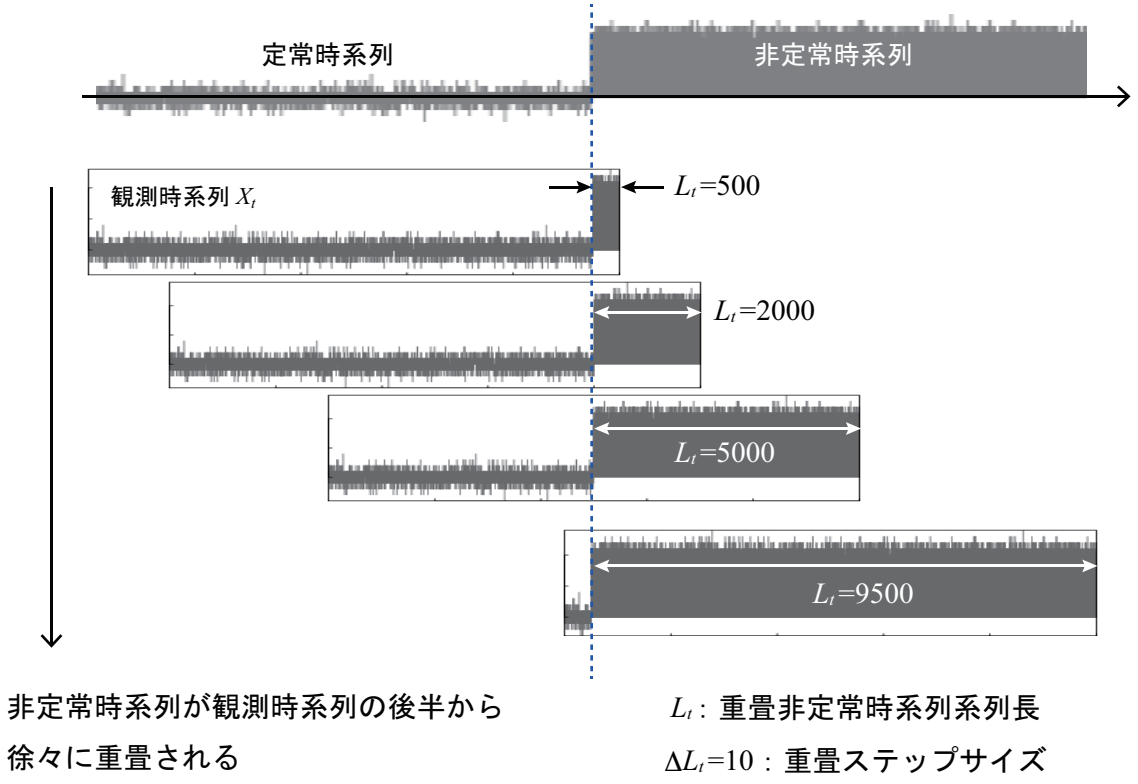


図 3.8 非定常時系列に対する時間特性シミュレーション

Akita University

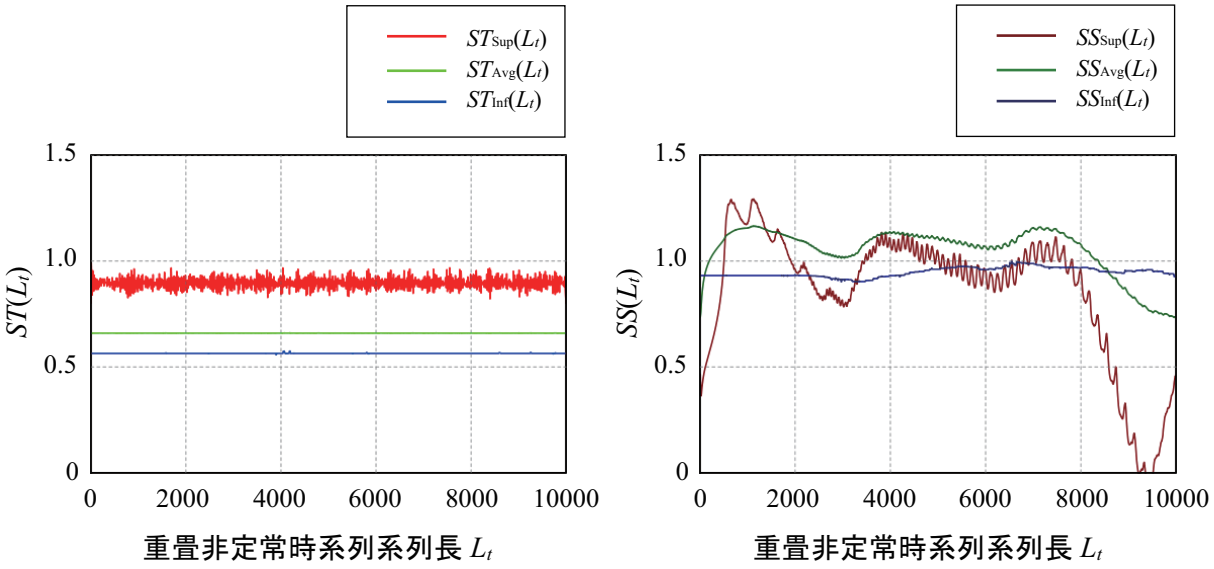
レベルシフト時系列に対する各 Slope 値の時間特性を図 3.9 に示す．図 3.9(a) は Slope 値 $ST(L_t)$，図 3.9(b) は Slope 値 $SS(L_t)$，図 3.9(c) は図 3.9(a) の重畳非定常時系列系列長 L_t の $0 \leq L_t \leq 1000$ の範囲，つまり非定常時系列重畳直後における Slope 値 $ST_{\mathrm{Sup}}(L_t)$ を示している．

図 3.9(a) より，$ST_{\mathrm{Avg}}(L_t)$ と $ST_{\mathrm{Inf}}(L_t)$ は L_t に対してほぼ変化を示さず，$ST_{\mathrm{Sup}}(L_t)$ は重畳直後から定常時の値から上昇することが観測された．具体的には図 3.9(c) より最初のステップで急激に約 1.0 まで上昇，次のステップで値が低下するが，その後振動を繰り返しながら非定常時系列が全体に重畳されるまで約 0.9 辺りで収束，振動を繰り返している．また，図 3.9(b) より，$SS_{\mathrm{Sup}}(L_t)$，$SS_{\mathrm{Avg}}(L_t)$ も重畳直後から値が上昇しているが，大きく変動を繰り返しながら，L_t が大きくなると逆に値が低下することが観測された．重畳直後の即応性および経時変化に対する安定性を考慮すると，ST_{Sup} がレベルシフト時系列の検知指標として有効であると推測される．

周期的時系列に対する各 Slope 値の時間特性を図 3.10 に示す．図 3.10(a)，(b)，(c) に示すグラフは前述のレベルシフト時系列のシミュレーションと同様の内容となっている．

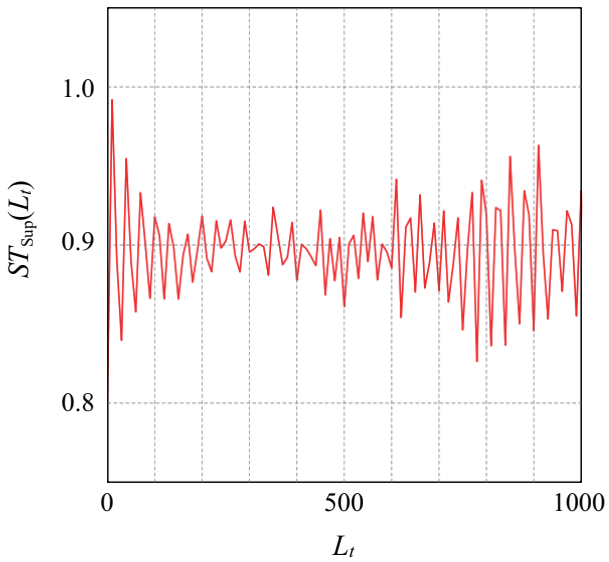
図 3.10(a) より，$ST_{\mathrm{Sup}}(L_t)$ および $ST_{\mathrm{Inf}}(L_t)$ はレベルシフト時系列に対する経時変化とほぼ同様となっているが，$ST_{\mathrm{Avg}}(L_t)$ は L_t に比例して上昇することが観測された．図 3.10(c) より，$ST_{\mathrm{Sup}}(L_t)$ においては，急激に上昇し，その後振動を繰り返すが，レベルシフト時系列と比べてより安定的に約 1.0 の値に収束することが観測された．このことから，周期的時系列においても ST_{Sup} が検知指標として有効であると推測される．また，図 3.10(b) より，各 SS は，大まかな変動傾向はレベルシフト時系列と同様となるが，$SS_{\mathrm{Avg}}(L_t)$ および $SS_{\mathrm{Inf}}(L_t)$ はレベルシフト時系列とは異なって値が減少し，それぞれ水平に近い値を示すことが観測された．つまり，SS_{Avg} および SS_{Inf} が周期性の存在を示す検知指標として有効であると推測される．

Akita University



(a) Slope of Thigh

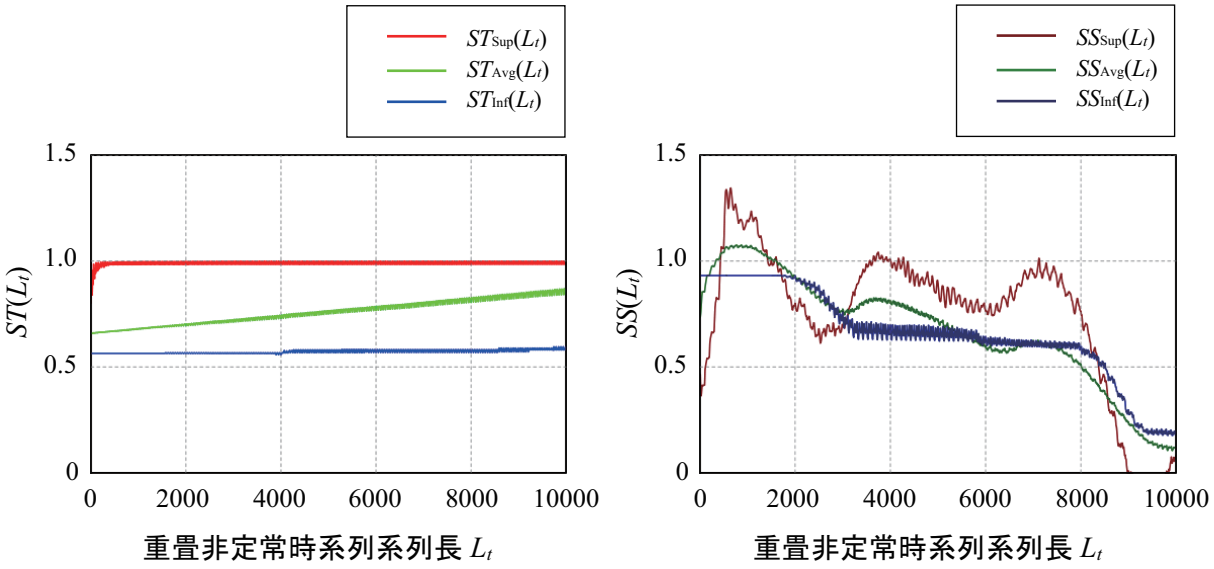
(b) Slope of Shin



(c) $0 \leqq L_t \leqq 1000$ における $ST_{Sup}(L_t)$

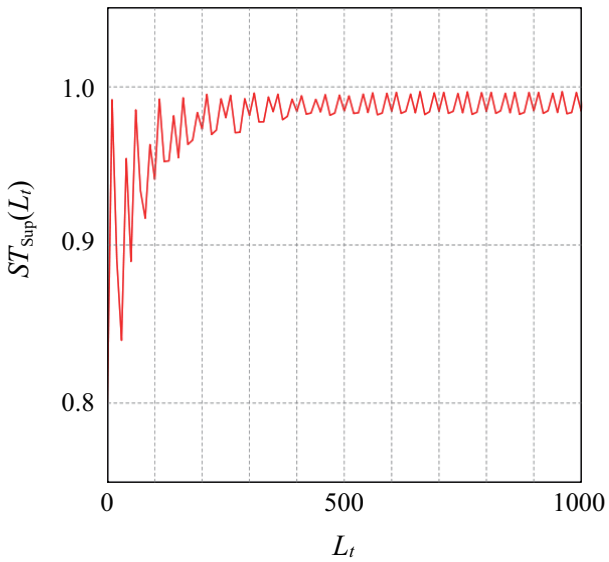
図 3.9 レベルシフト時系列に対する時間特性

Akita University



(a) Slope of Thigh

(b) Slope of Shin



(c) $0 \leqq L_t \leqq 1000$ における $ST_{Sup}(L_t)$

図 3.10 周期的時系列に対する時間特性

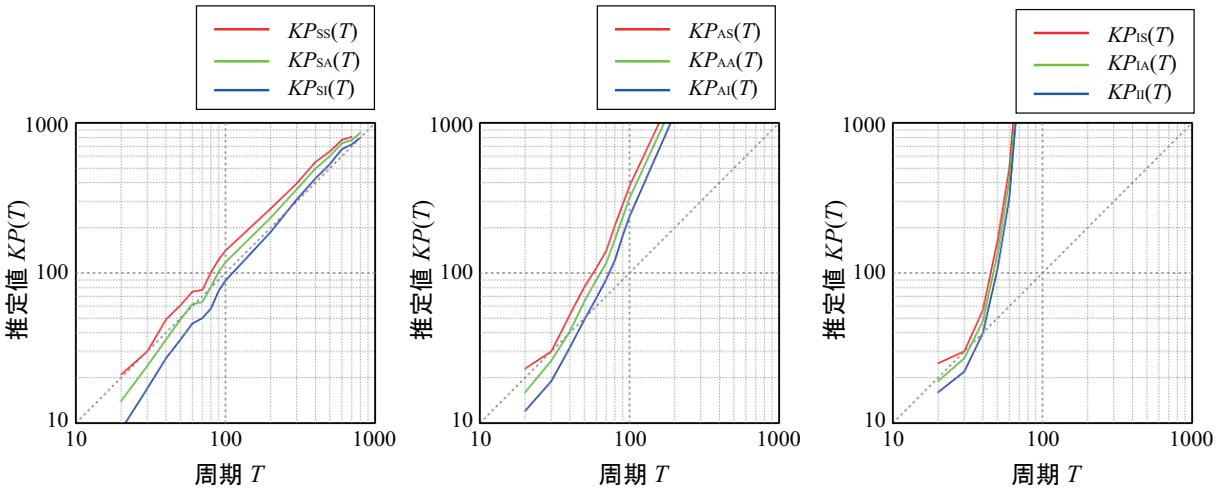
Akita University

3.5 周期推定

3.3で提案した周期推定法の有効性を評価するため，振幅 $P = 10$，デューティ比 $D = 50\%$，周期 T を $10 \leq T \leq 1000$ と変化させたシミュレーション時系列に対して，周期推定実験を試みた．

実験より算出された KP を図3.11に示す．図3.11(a)は導出範囲RTにおいて ST_{Sup} を選択したときの推定値 KP_{SS}, KP_{SA}, KP_{SI}, 図3.11(b)は導出範囲RTにおいて ST_{Avg} を選択したときの推定値 KP_{AS}，KP_{AA}，KP_{AI}，図3.11(c)は導出範囲RTにおいて ST_{Inf} を選択したときの推定周期 KP_{IS}，KP_{IA}，KP_{II} を示している．各グラフの横軸はシミュレーション時系列の周期 T，縦軸は推定値 KP を表し，右上がりの破線で示される対角線は，推定周期の理想直線を表す．プロット点の存在しない周期 T は，提案手法の制約条件によって算出されなかったものである．

結果より，導出範囲RTにおいてSlope値 ST_{Sup} を選択した場合，導出範囲RSのどのSlope値においても理想直線に近似した形となり，良好な周期推定が行われている．ST_{Avg}，ST_{Inf} を選択した場合は，短周期では推定できているが，長周期になるにつれて誤差が大きくなることが観測された．これより，周期推定に対しては ST_{Sup} を選択するのが有効であることが示され，本手法により目標とした周期推定範囲を概ね推定できることが示された．



(a) ST_{Sup} を選択　(b) ST_{Avg} を選択　(b) ST_{Inf} を選択

図 3.11 周期推定結果

Akita University

3.6 まとめ

本章では，R/S Pox Diagram の特徴的プロット形状を定量的に表現する R/S Pox レッグライン特性を新たに定義して，本特性の特徴量を用いた周期推定法を提案した．また，異常トラフィックを想定した非定常性を有するシミュレーション時系列に対する特徴量の影響，ならびに周期推定を試みることで性能を評価し，検討を加えた．得られた成果をまとめると次のようになる．

- 非定常性に対する特徴量
  (1) レベルシフト時系列に対する上部プロット点群は，ST_{Sup}，SS_{Sup} で定量化でき，ST_{Sup} においては，振幅が大きくなると約 1.0 と安定値を示すことを明らかにした．
  (2) 周期的時系列に対する折れ曲がるプロット形状は，前・後半の傾きをそれぞれ ST と SS で定量化できることを明らかにした．
  (3) 周期性に対する特徴量は，振幅，デューティ比の変動に対してロバスト性を有することを明らかにした．
  (4) 時間特性より，非定常性のトラフィック変化点が時系列 X_t に存在している間，ST_{Sup} は定常時より高い値を示すことを明らかにした．

- 周期推定
  (1) 周期的時系列の周期 T は，本特性の KP で推定可能であることを明らかにした．
  (2) 導出範囲 RT において ST_{Sup} を選択した場合，他に比べ広範囲で周期推定が可能であることを明らかにした．

参考文献

[1] 高橋秋典，五十嵐隆治，上田浩，岩谷幸雄，木下哲男：R/S Pox レッグライン特性，情報処理，Vol.54，No.6，pp.1761-1770 (2013).

[2] 独立行政法人情報処理推進機構：高トラフィック観測・分析法に関する技術調査，pp.42-43 (2004)

[3] 角田裕，荒井健二郎，和泉勇治，根元義章：パルス型DoS攻撃の被害軽減のためのトランスポート層プロトコルの通信制御に関する検討，電子情報通信学科技術研究報告，Vol.107，No.18，pp.43-48 (2007).

[4] 三鑰達真，吉田和幸：長期的スキャンニングを対象としたスキャン攻撃検知システム，電子情報通信学会技術研究報告，Vol.107，No.449，pp.39-44 (2008).

第4章　実トラフィックデータによるポートスキャン検知実験

4.1　はじめに

本論文で観測対象とした長期的ポートスキャン攻撃[1]について説明する．長期的ポートスキャン攻撃とは，悪意ある攻撃者が対象とするホストの脆弱性を事前調査するために行うもので，使用する調査パケットを短時間で大量に送信せずに，時間間隔を置いて間欠的に少量の調査パケットを送信する攻撃である．このようにトラフィック量を抑制することで，しきい値を用いた検知法では検知しにくくなってしまう．また，この攻撃を検知するためにしきい値を低く設定すると，正常な通信も誤検知してしまう可能性があったり，パルス列の発生の度にしきい値を超え，アラートが多発してしまう可能性も懸念されるため，セキュリティー監視上，問題点が多く，その検知は難しい．

長期的ポートスキャン攻撃のトラフィックモデルを図4.1に示す．Tは攻撃トラフィックのパルス間隔，τはパルス幅，Pはパルスのピークレートを表し，Nは非攻撃時の定常パケットレートを表す．図のように，攻撃パケットが間欠的に送信されるため周期的時系列の特徴を有しており，提案手法による周期性の評価を用いることで，攻撃検知が可能と推測される．

そこで，R/S Poxレッグライン特性の実用性に対する評価として，実際のネットワークより観測されたトラフィックデータを用い，長期的ポートスキャントラフィックに対して以下に示す実験を試みた．

(1) ポートスキャントラフィックに対する周期推定
(2) 特徴量経時変化によるポートスキャン検知性能

Akita University

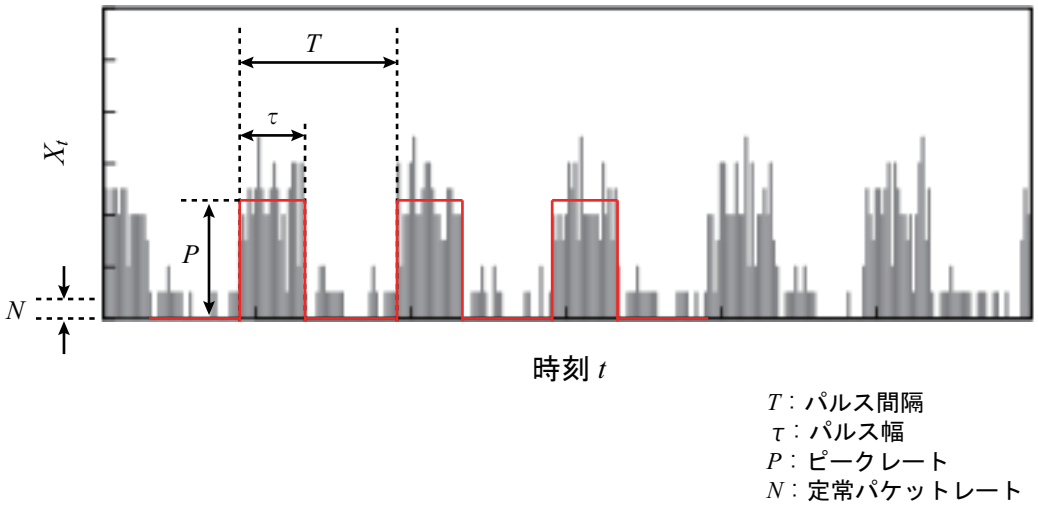


図 4.1 長期的ポートスキャンモデル

Akita University

4.2 ポートスキャントラフィックに対する周期推定

4.2.1 実験概要

周期推定実験の概要について説明する．まず，実環境から tcpdump[2] を用いて取得されたトラフィックデータより，パケットデータ解析ツールの Wireshark[3] を用いてヘッダを解析し，長期的ポートスキャントラフィックが観測されるデータを選択する．このトラフィックデータから，学外から到着した TCP SYN パケットのみをカウントした時系列データを生成し，長期的ポートスキャントラフィックが全体に反映されている導出時系列 X_t を抽出する．このとき，この時系列 X_t には，攻撃トラフィック以外にも正常な通信で発生した TCP SYN パケットも含まれることになる．ここで，時系列 X_t の測定単位時間 Δt は 0.02s，系列長 N は 3000 点，つまり解析に用いたデータは 60s 間に計測されたものとなる．この時系列データに対して周期推定実験を試みた．

4.2.2 トラフィックデータの詳細

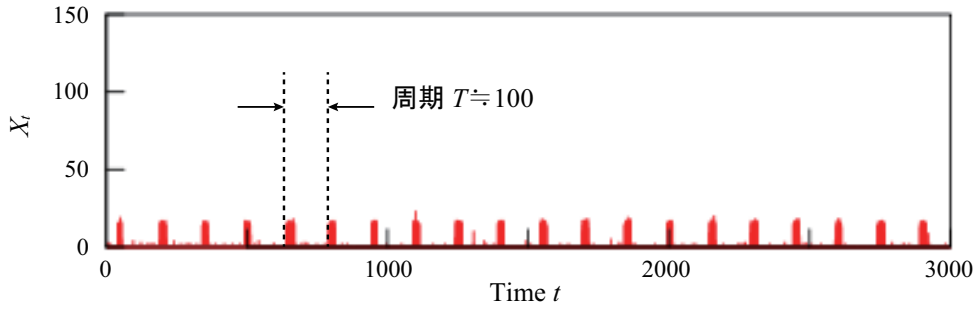
実験に用いたポートスキャン攻撃の詳細を表 4.1 に，そのときの時系列 X_t のグラフを図 4.2 に示す．取得されたパケットデータ内に発生していた長期的ポートスキャン攻撃のうち，本実験では表 4.1 に示される 3 つのスキャン攻撃に対して実験を行った．それぞれ Scan1，Scan2，Scan3 と呼ぶことにする．

Scan1 はポート 4899(TCP) に対するポートスキャンで，Microsoft Windows[4] ベースのリモート操作を行う RAdmin[5] の脆弱性を狙ったものである．Scan2 は，一般的に Web サーバアクセスに使用されるポート 80(TCP) に対するもので，公開サーバゆえセキュリティ上の脆弱性が多く，CodeRed[6] や Nimda[7] といった不正アクセスのセキュリティーホールとなりやすい．Scan3 は Telnet で使用されるポート 23(TCP) に対するもので，脆弱性が存在した場合，システムへの侵入を許してしまい，様々なサービスへの影響が懸念される．表 4.1 の周期 T，パルス幅 τ，振幅 P は，それぞれの攻撃の時系列 X_t において，目視により計測した値である．Scan1 の周期的時系列の特徴として，周期 T，振幅 P は比較的安定的な値だが，パルス幅 τ は約 10〜20 と変動が生じている．Scan2 では，周期 T，パルス幅 τ は安定的だが，振幅 P は約 30〜90 と変動が生じている．Scan3 は，各パラメータとも比較的安定した値を示している．

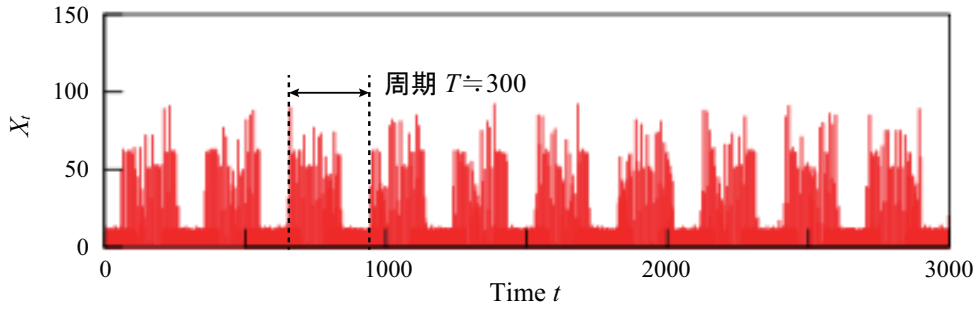
Akita University

表 4.1 ポートスキャン攻撃の詳細

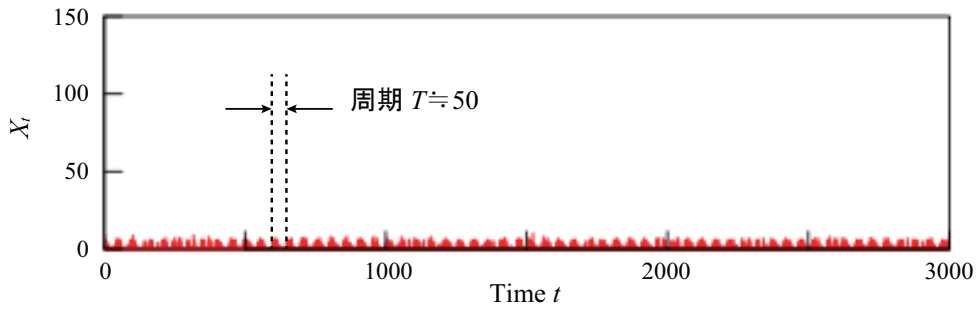
	Scan1	Scan2	Scan3
計測日	2008/08/26	2008/08/27	2008/08/30
宛先ポート	4899	80	23
周期 T	約 150	約 300	約 50
パルス幅 τ	約 10〜20	約 200	約 20
振幅 P	約 15	約 30〜90	約 5



(a) Scan1 の時系列 X_t



(b) Scan2 の時系列 X_t



(c) Scan3 の時系列 X_t

図 4.2 長期的ポートスキャントラフィック時系列 X_t

Akita University

4.2.3 実験結果

周期推定実験により算出されたKnee Pointの値を表4.2に示す．また，そのときに導出されたR/S Pox Diagramを図4.3に示す．KPを算出するための導出範囲RTの傾きSTは，3.5で示したように良好な推定結果が得られたST_{Sup}を用いて，KP_{SS}，KP_{SA}，KP_{SI}の3点を算出し結果を比較した．図4.3に示されている緑の破線は，各攻撃の実測周期Tと同じ値の任意長区間の区間長$n = T$を表し，赤の破線がKP_{SS}，黒の破線はそれぞれKP_{SA}，KP_{SI}を表している．

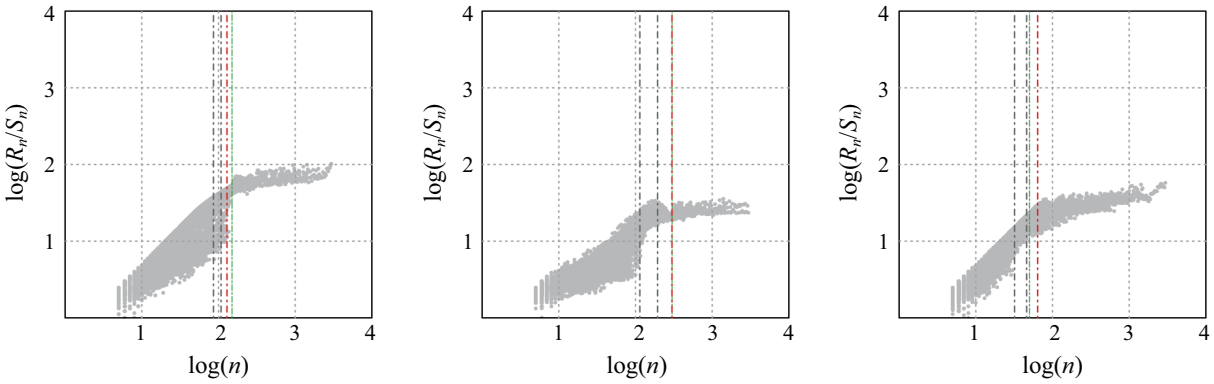
図4.3より，各スキャン攻撃に対するR/S Pox Diagramは，その周期的特徴により折れ曲がるプロット形状を確認できた．また，図4.3(a)および図4.3(c)においては，プロット形状の折れ曲がる部分と周期Tと等しい区間長nがほぼ一致している．しかし，図4.3(b)のScan2の場合は，折れ曲がる部分が$\log(n) = 2$付近のように観測でき，区間長$n = T$とはずれているように観測される．ここで，図4.2(b)のScan2の時系列X_tを確認すると，振幅$P = 50$程度で確認できる周期$T = 300$のパルス時系列の他に，振幅$P = 10$程度で断続的に到着するパケット時系列が存在している．R/S Pox Diagramの$\log(n) < 2$の範囲においても，下限点群の傾きST_{Inf}も短周期のパルス時系列のために傾きが小さくなっているため，折れ曲がる部分がずれたように観測されたと推測される．このことから，R/S Pox Diagramには複数の異なる周期性が混在した場合，周期的特徴は周期成分の数だけ顕現すると考えられ，その場合，本提案手法である周期推定法では最も大きい周期成分を推定する性能を有していると考えられる．

表4.2の推定結果を比較すると，全てのスキャン攻撃に対して，KP_{SS}，KP_{SA}，KP_{SI}のうち，各攻撃の実測周期Tの値により近い値を示したのはKP_{SS}であった．また，Scan1のようにパルス幅τが変動している時系列X_tや，Scan2のように振幅Pが変動している時系列X_tの場合でも，良好な周期推定を行っていることが確認できた．つまり，提案する周期推定法はパルス幅τ，振幅Pの変動に対しロバスト性を有していると推測できる．

Akita University

表 4.2 周期推定結果

	Scan1	Scan2	Scan3
実測周期 T	約 150	約 300	約 50
KP_{SS}	129	299	64
KP_{SA}	108	194	46
KP_{SI}	86	114	32



(a) Scan 1　(b) Scan 2　(c) Scan 3

図 4.3 各攻撃時系列 X_t における R/S Pox Diagram

Akita University

4.3 特徴量経時変化によるポートスキャン検知性能

4.3.1 実験概要

R/S Pox レッグライン特性の特徴量はレベルシフトや周期性といった非定常性に対して反応を示すことから，特徴量の経時変化を観測することで異常検知が可能と推測される．そこで, 実運用でのトラフィック異常検知を想定したシミュレーションとして，実トラフィックにおける長期的ポートスキャン攻撃の検知実験を行い，検討を加えた．

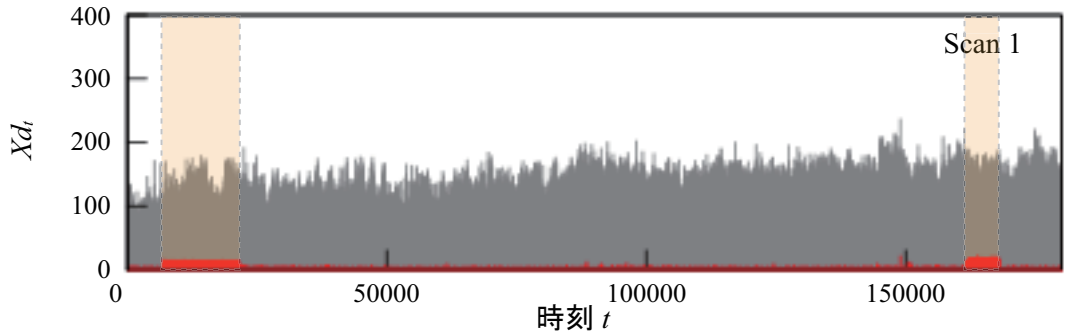
特徴量経時変化の導出手順について説明する．実環境から測定単位時間 Δt で計測された系列長 Nd の実験用時系列を Xd_t，特徴量導出に用いる系列長 N の観測時系列を X_t とする．まず，実験用時系列 Xd_t の $0 \leq t < N$ の系列を観測時系列 X_t として特徴量を導出する．その後，時刻 t を Δs だけシフトさせて，つまり，Xd_t の範囲 $\Delta s \leq t < N + \Delta s$ を観測時系列 X_t として特徴量を導出し，Xd_t の最後まで繰り返すことで経時変化を導出する．このとき，シフト量 Δs は，実時間上における特徴量導出時間間隔 $\Delta t \times \Delta s$ 秒を表す値である．

本実験では，4.2 と同様に測定単位時間 Δt を 0.02s，観測時系列 X_t の系列長 N を 3000 として，1 分間のデータを用いて特徴量を導出する．シフト量 Δs は，実運用上リアルタイム性を損なわない時間間隔を 1 秒間と想定し，$\Delta s = 1s/0.02s = 50$ と設定した．

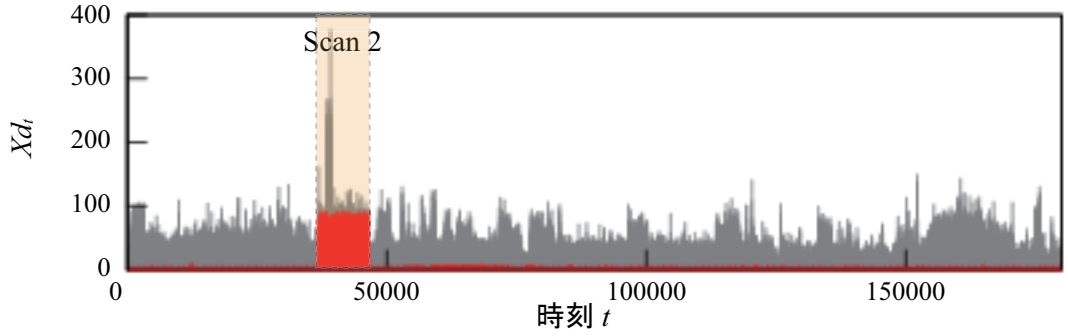
4.3.2 トラフィックデータの詳細

実験に用いた長期的ポートスキャン攻撃が含まれているトラフィック時系列データを図 4.4 に示す．図 4.4(a) は前項の周期推定実験で用いたポートスキャン攻撃 Scan1 が含まれている 2008 年 8 月 26 日 9:00〜10:00 の 1 時間に計測されたトラフィック時系列, 図 4.4(b) は Scan2 が含まれている 2008 年 8 月 27 日 3:00〜4:00 のトラフィック時系列, 図 4.4(c) は Scan3 が含まれている 2008 年 8 月 30 日 11:00〜12:00 のトラフィック時系列を表しており，各ポートスキャン攻撃が発生している範囲を肌色で表している．灰色で示されているデータは秋田大学キャンパスネットワークの学外より到着した全てのパケットの時系列データ，赤色で示されているデータは，その中の TCP SYN パケットのみから作成した時系列データである．本実験では, この時系列を対象実験用時系列 Xd_t とした．それぞれの時系列データを Data1, Data2，Data3 と呼ぶことにする．ここで，Data1 には Scan1 の他にもポートスキャン攻撃が発生しているため，本実験ではその攻撃も検知対象とした．

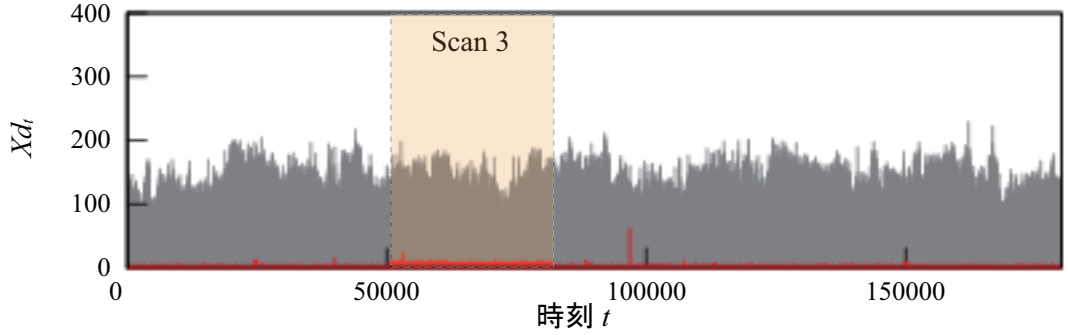
Akita University



(a) Data 1 の時系列 Xd_t



(b) Data 2 の時系列 Xd_t



(c) Data 3 の時系列 Xd_t

図 4.4 ポートスキャントラフィックの実験用時系列 Xd_t

Akita University

4.3.3 実験結果

まず，特徴量Slope値の経時変化について検討する．

Data1に対する各Slope値ST，SSの経時変化グラフを図4.5，Data2に対する各Slope値の経時変化グラフを図4.6，Data3に対する各Slope値の経時変化グラフを図4.7に示す．各グラフの時刻tは，1時間の実験用時系列Xd_tに対して観測時系列X_tを1秒間ずつ導出したので，3600ポイント導出されたことになる．また，肌色で示されている領域はポートスキャン攻撃が発生している範囲を表している．Slope値SSの各グラフにおいて，周期性判定として設定したしきい値γを赤の破線で表している．

図4.5，図4.7より，Scan1，Scan3に対するSlope値は同様の変動傾向を示した．具体的には，ポートスキャン発生直後にST_{Sup}，ST_{Avg}，ST_{Inf}は定常時に比べ値が増加し，ポートスキャン終了後，しばらくしてから定常時の値に戻ることが観測された．また，SS_{Sup}，SS_{Avg}，SS_{Inf}は，定常時に激しい変動を呈しているが，ポートスキャン攻撃時にはSSにおいて判定しきい値γを下回ることが観測された．しかし，SS_{Sup}では定常時にも頻繁にしきい値γを下回っており，ここからSS_{Sup}には周期性以外のトラフィック事象に対しても反応を示すことがわかる．これは，3.4.5の図3.9(b)で示したように，突発的にトラフィック量が増加・減少をするレベルシフト時系列に対しては，観測時系列X_t内のトラフィック変化点の観測位置によってSS_{Sup}は値が減少することが要因と考えられる．これに対して，SS_{Avg}，SS_{Inf}はポートスキャン発生時のみしきい値γを下回ることから，周期性の判定指標として有効であると推測される．

図4.6より，Scan2に対しては，Slope値STにおいてScan1およびScan3とは異なる変動傾向が観測された．具体的には，ST_{Sup}はポートスキャン発生直後，値は増加するが，観測時系列X_t全体にポートスキャントラフィックが反映されると値は定常時より小さくなることが観測された．また，ST_{Avg}，ST_{Inf}は発生直後より値が減少することが観測された．この要因として，4.2.2で述べたように，Scan2では複数の周期成分が混在して，短周期成分に対する反応が影響していると推測される．このように，ポートスキャントラフィックの時系列パターンによりSTの変動傾向は確定しないが，ポートスキャン発生直後のST_{Sup}は増加することが確認できた．ここから，ST_{Sup}は異常トラフィック発生時の検知指標として有効であると推測される．

Akita University

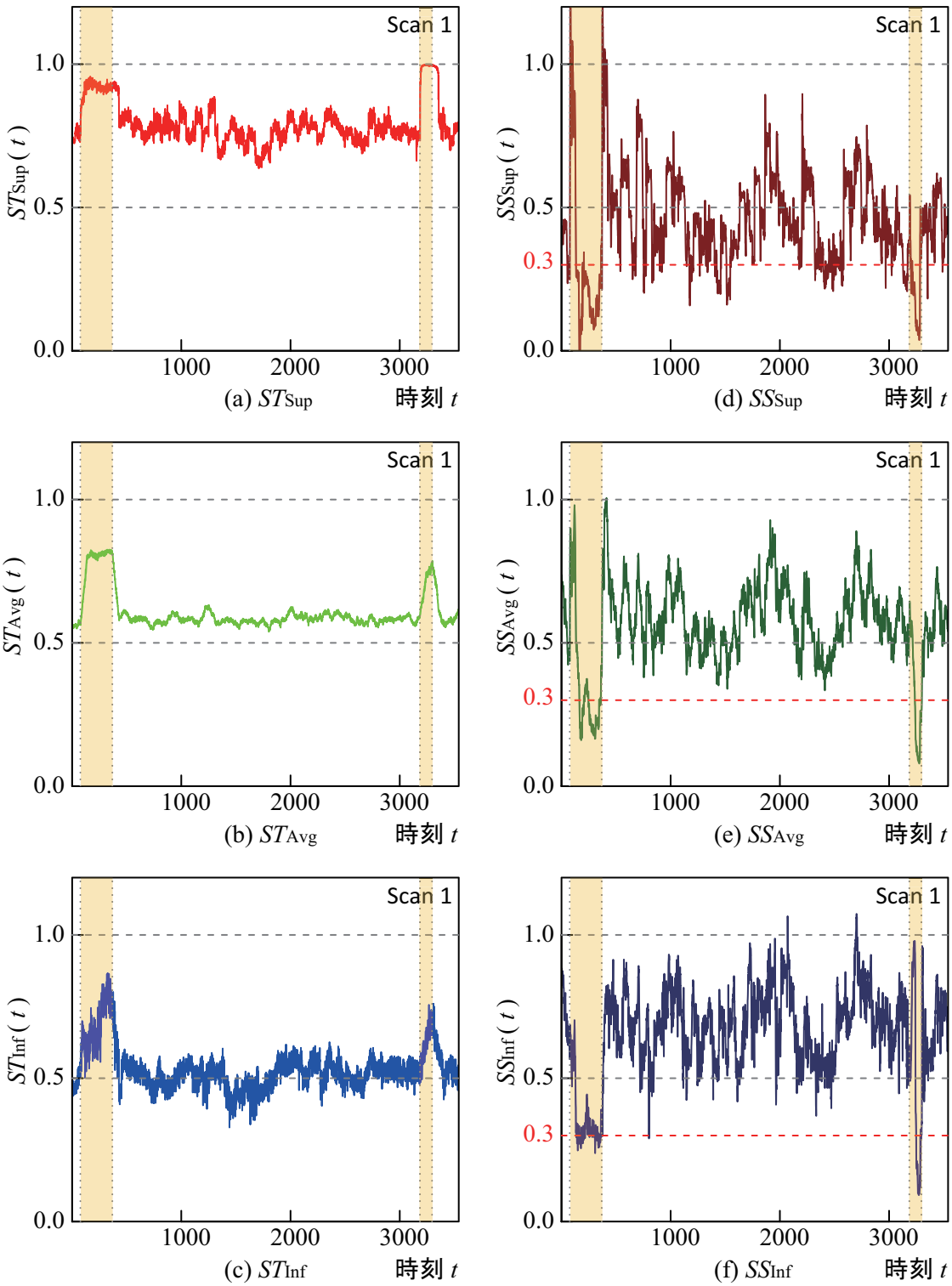


図 4.5 Data1 に対する Slope 値の経時変化

Akita University

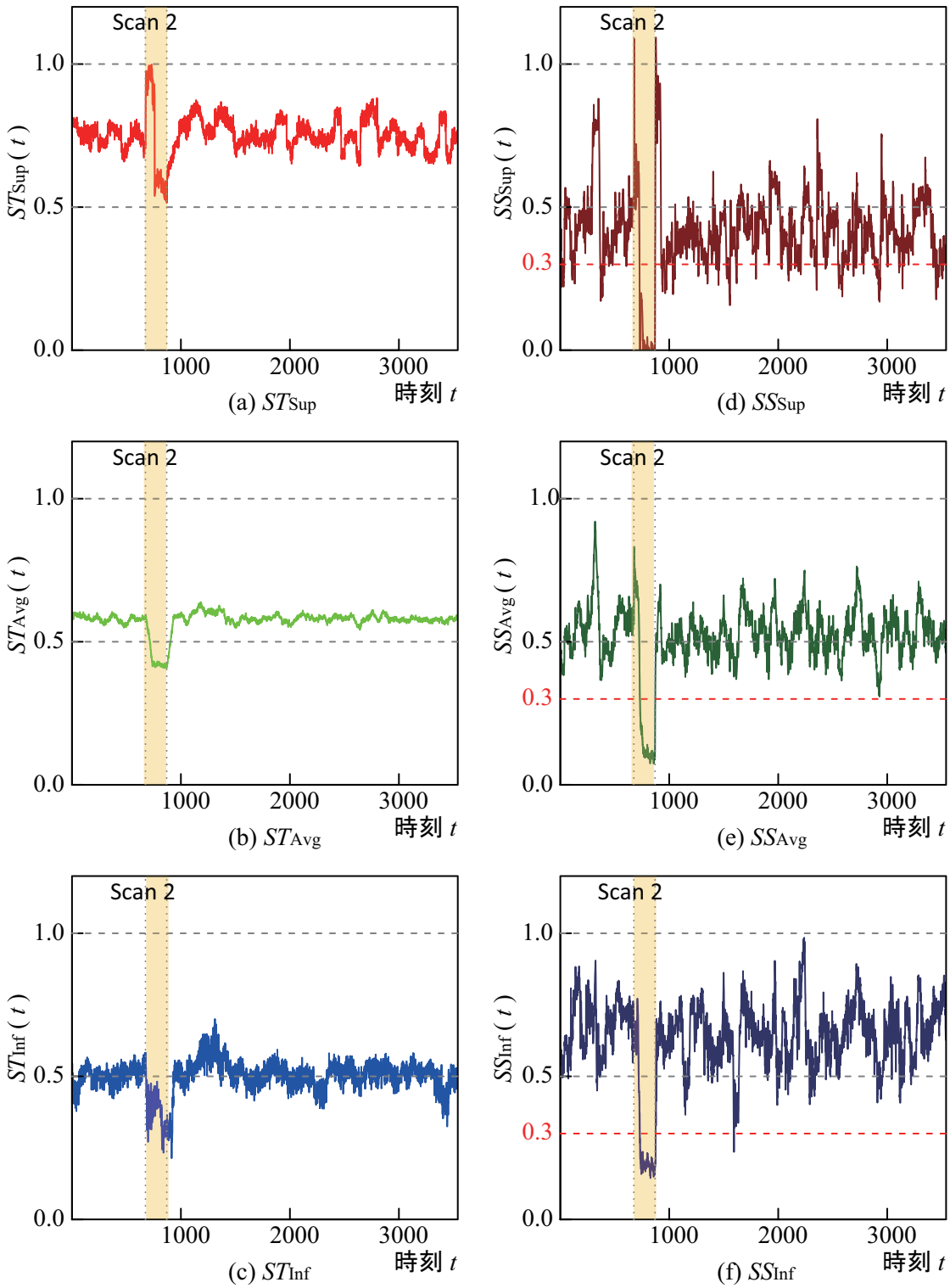


図 4.6 Data2 に対する Slope 値の経時変化

Akita University

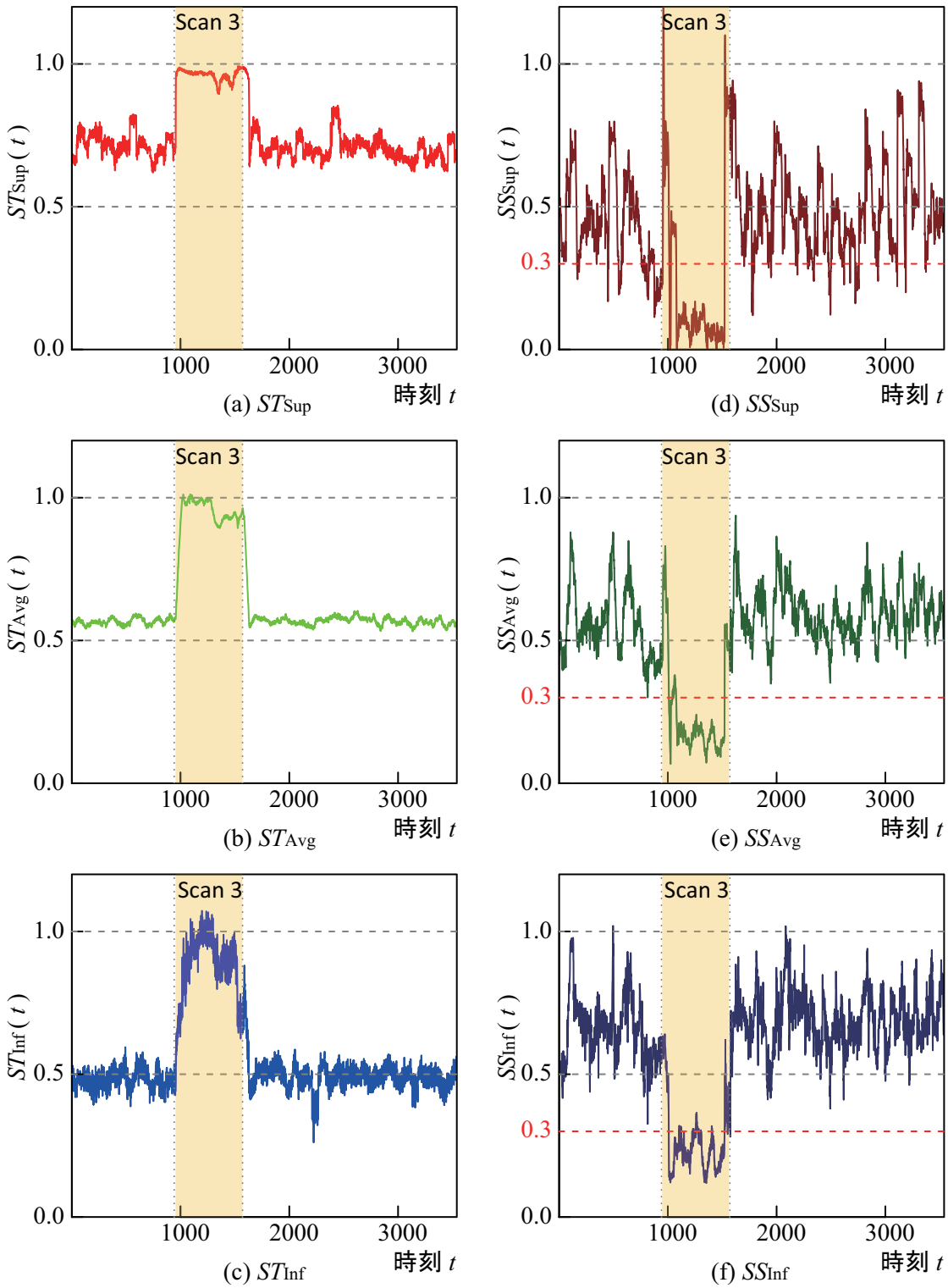


図 4.7 Data3 に対する Slope 値の経時変化

Akita University

次に推定周期KPの経時変化について検討する．

Data1に対する推定周期KPの経時変化グラフを図4.8，Data2に対する経時変化グラフを図4.9，Data3に対する経時変化グラフを図4.10に示す．各グラフの時刻tは，Slope値の経時変化グラフと同様に3600ポイントとなっている．また，肌色で示されている領域はポートスキャン攻撃が発生している範囲を表している．赤の破線は，それぞれScan1，Scan2，Scan3の実測周期を表している．

図4.8，図4.9，図4.10より，ポートスキャン攻撃が継続している間に，推定周期KP_{SS}，KP_{SA}，KP_{SI}が導出されることが観測できた．Data1，Data2においてはKP_{SS}が，Data3においてはKP_{SA}が良好な推定結果を示したが，KP_{SS}は攻撃以外の部分においても導出されており，誤推定が行われている．これは，SS_{Sup}がレベルシフト時系列で周期判定しきい値γを下回ることに起因している．つまり，KP_{Sup}は長期的ポートスキャンの検知指標としては有効ではないが，推定周期は良好な結果を示すことから，実用化においてはポートスキャン検知指標としてはKP_{Avg}を利用して，検知判定後，KP_{Sup}より周期を推定するといった運用手順の検討も有効と推測される．

Akita University

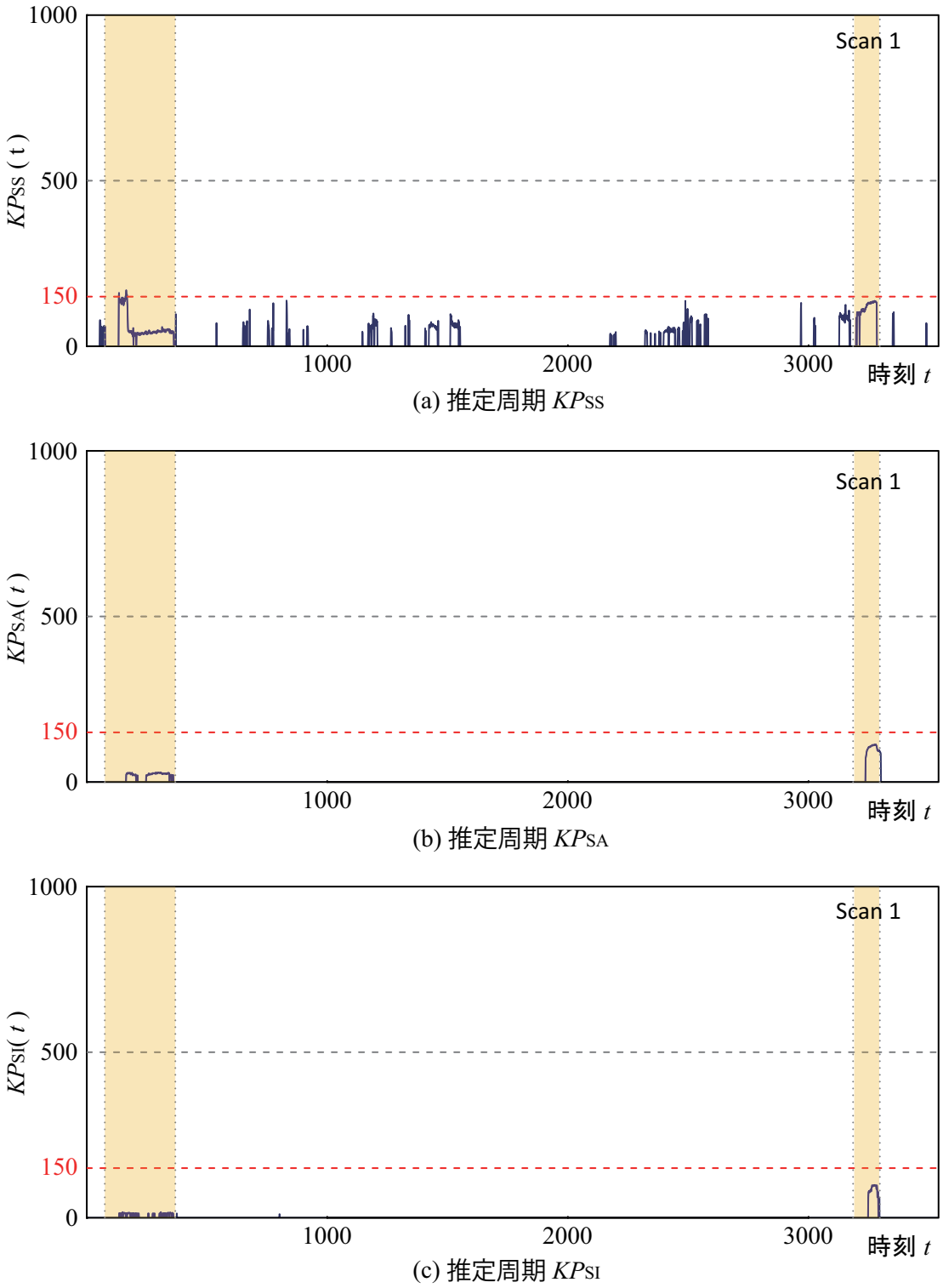
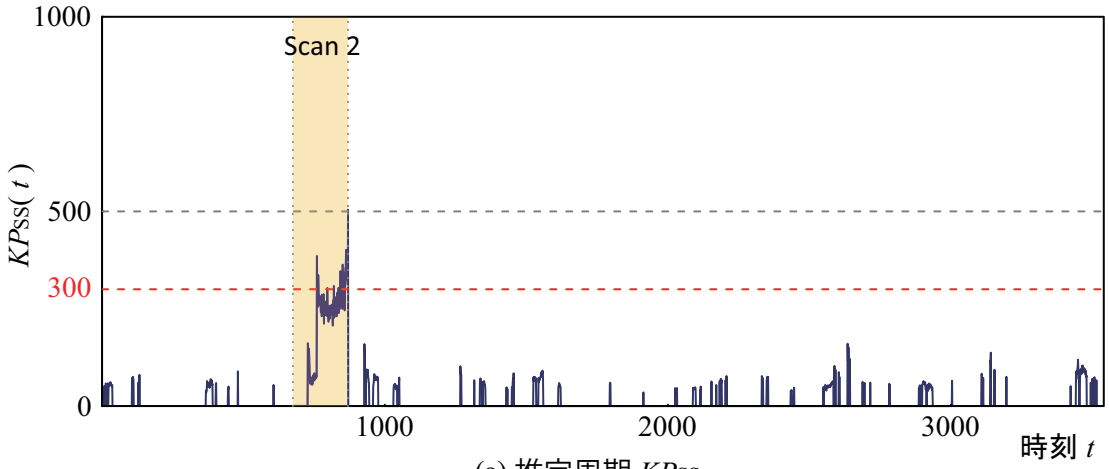
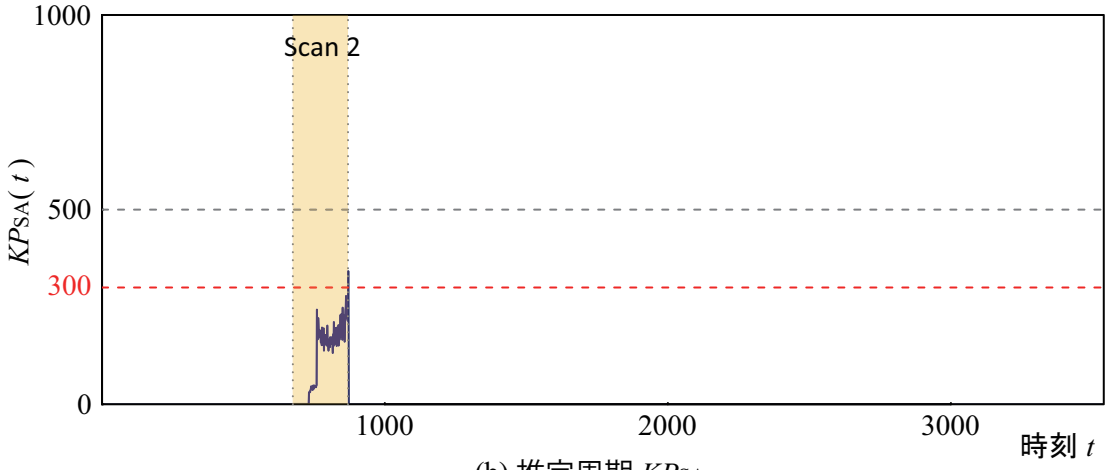


図 4.8 Data1 に対する KP の経時変化

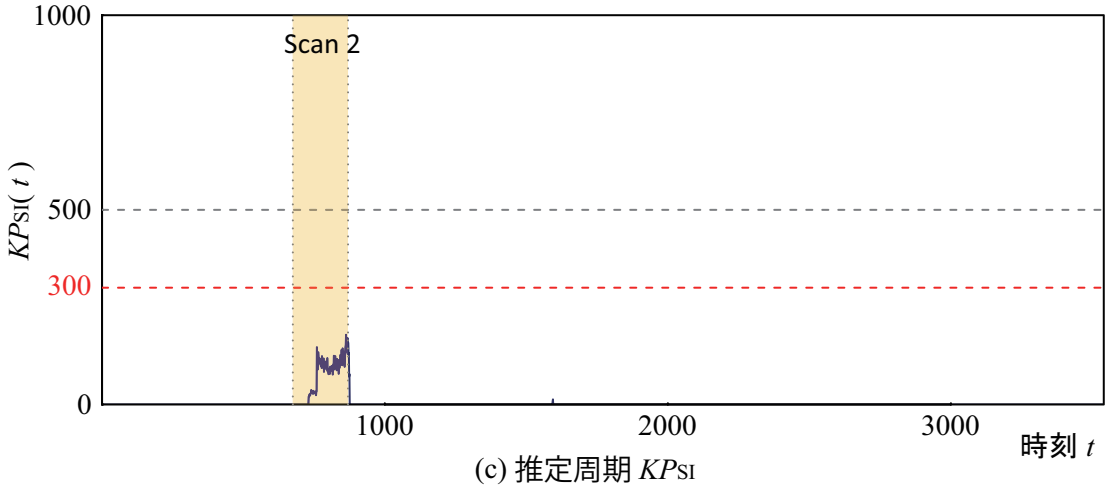
Akita University



(a) 推定周期 KP_{SS}



(b) 推定周期 KP_{SA}



(c) 推定周期 KP_{SI}

図 4.9 Data2 に対する KP の経時変化

Akita University

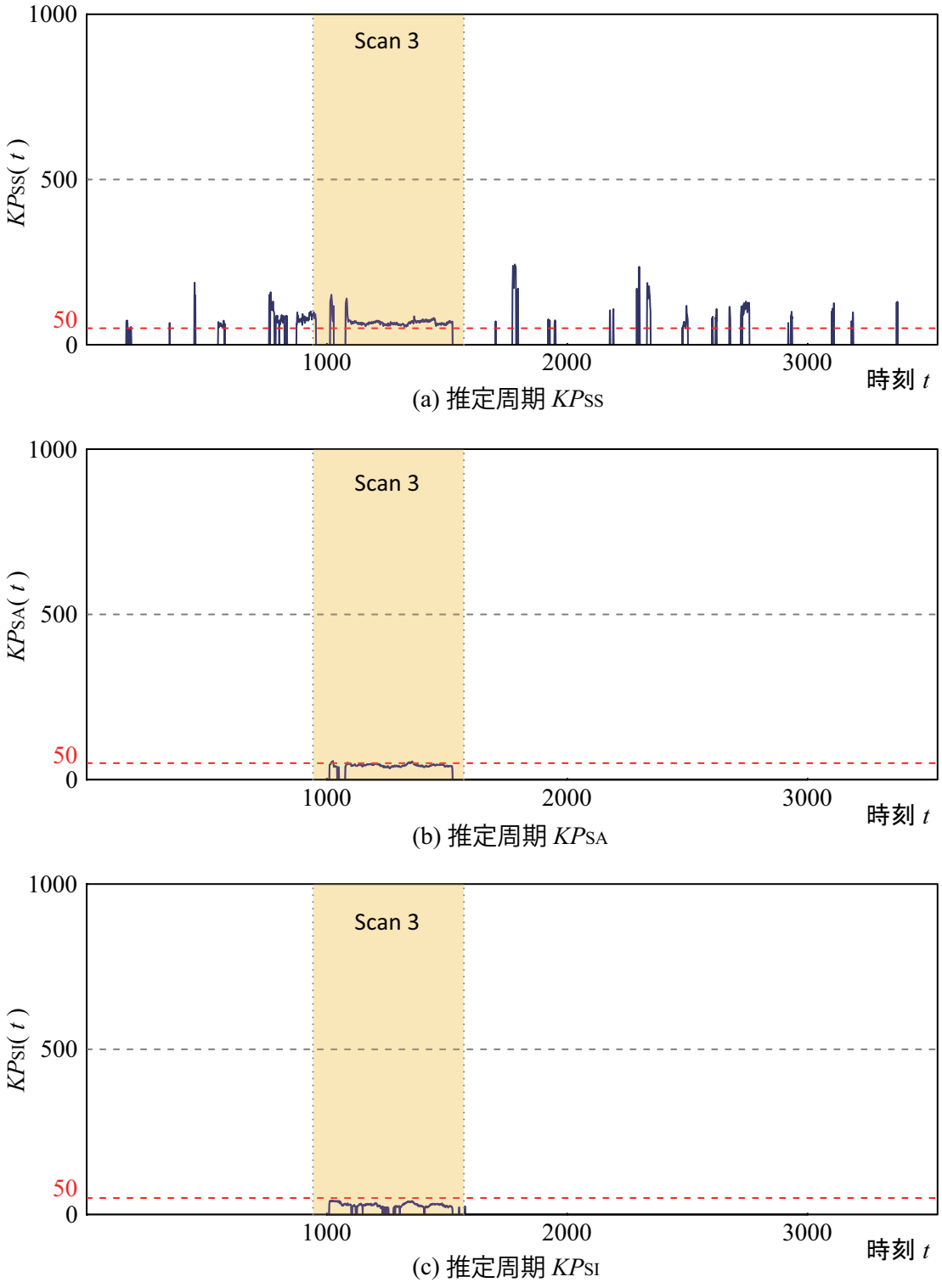


図 4.10 Data3 に対する KP の経時変化

Akita University

4.4 まとめ

本章では，R/S Pox レッグライン特性の実用性に対する評価として，秋田大学キャンパスネットワークより取得された実トラフィックデータを用いて，長期的ポートスキャントラフィックに対する検知実験を試み，検討を加えた．得られた成果をまとめると次のようになる．

(1) 実環境で発生した長期的ポートスキャンに対する周期推定は，KP_{SS} が他の値に比べ，実測周期に近い値を推定できることを明らかにした．
(2) パルス幅 τ や振幅 P に変動がある場合でも，周期推定が可能であるというロバスト性を明らかにした．
(3) SS_{Sup} は時間特性によりレベルシフト時系列に対しても周期判定しきい値 γ を下回るため，KP_{SS} による周期推定は誤推定をしてしまうことを明らかにした．
(4) KP_{SA} は周期性にのみ反応することから，長期的ポートスキャントラフィックの検知指標として有効であることを明らかにした．

参考文献

[1] 三鑰達真, 吉田和幸：長期的スキャンニングを対象としたスキャン攻撃検知システム，電子情報通信学会技術研究報告，Vol.107，No.449，pp.39-44 (2008).

[2] TCPDUMP&LiBPCAP, http://www.tcpdump.org/

[3] Wireshark, http://www.wireshark.org/

[4] Microsoft Windows, http://windows.microsoft.com/

[5] Radmin: Remote Control Software, http://www.radmin.com/

[6] CERT Advisory CA-2001-19: "Code Red" Worm Exploiting Buffer Overflow In IIS Indexing Service DLL, http://www.cert.org/historical/advisories/CA-2001-19.cfm

[7] CERT Advisory CA-2001-26: Nimda Worm, http://www.cert.org/historical/advisories/ca-2001-26.cfm

Akita University

第5章　結論

本論文では，R/S 解析法から得られる R/S Pox Diagram に顕現する非定常的時系列に対する特徴的プロット形状のトラフィック事象判別への積極的応用を検討し，非定常性を定量化する新たな特性である R/S Pox レッグライン特性を提案した．さらに，実用に対する有効性について，ネットワーク脅威となる長期的ポートスキャン攻撃への検知性能について検証した．以下に本論文で得られた主な知見をまとめ，これらの工学的意義についてまとめる．

5.1　本論文により得られた知見

第 1 章では，インターネットの現状，およびネットワークトラフィック異常検知の重要性について指摘し，バースト性を有するトラフィック時系列の自己相似性に着目した既存研究について概観した．ここから，R/S 解析法に顕現する非定常性に対する特徴を定量化する新たな特性を提案するという目的を述べるとともに，本論文の内容について述べた．

第 2 章では，R/S 解析法に対する考察を行い，トラフィック事象変化に対する即応性の問題点を改善する手法を検討した．さらに，非定常性に対する R/S Pox Diagram の特徴的プロット形状の発生要因についてシミュレーションを行い検討したところ，次のような結果が得られた．

(1) 本論文で提案した反転時系列を用いた解析手順により,「直近」で発生するトラフィック事象変化が未計算区間に影響されずに全任意長区間に反映されるため，即応性が高まることを明らかにした．
(2) R/S Pox Diagram の上部に発生するプロット点群は，突発的トラフィック増減を表現するレベルシフト時系列により発生し，折れ曲がるプロット形状は，時間間隔を置いて到着する周期的時系列により発生することを明らかにした．
(3) 折れ曲がるプロット形状の変曲点を示す任意長区間の区間長 n は，周期的時系列の周期 T と同等となることを明らかにした．

第 3 章では，第 2 章で明らかにした R/S Pox Diagram の特徴的プロット形状を定量化する新たな特性，および周期的時系列に対する周期推定法を提案し，シミュ

Akita University

レーション時系列を用いてその性能について検討を加えたところ，次のような結果が得られた．

(1) 特徴的プロット形状は，折れ曲がる特徴点を境に前半・後半のプロット点群の傾きを定量化することで表現できることを明らかにした．
(2) 非定常性として検証したレベルシフト時系列や周期的時系列のトラフィック変化点に対しては，ST_{Sup} が定常時に比べ高い値を示し，周期性に対しては SS が水平方向に傾くため低い値を示すことを明らかにした．
(3) 提案特性は，振幅やデューティ比の変化に対してロバスト性を有することを明らかにした．
(4) 周期的時系列の周期 T は，ST と SS，それぞれの回帰直線の交点である KP より推定可能であることを明らかにした．

第 4 章では，周期推定法の実用性に対する評価として，実環境で観測された長期的ポートスキャントラフィックに対する周期推定，および特徴量の経時変化による検知実験を試み検討を加えたところ，次のような結果が得られた．

(1) 長期的ポートスキャントラフィックに対しては，KP_{Sup} が他の値に比べ実測周期に近い値を推定できることを明らかにした．
(2) SS による周期性判定では，SS_{Sup} がレベルシフトに対しても反応してしまうことから，KP_{SS} よりも KP_{SA} のほうがポートスキャン検知指標として有効であることを明らかにした．

5.2 本論文の工学的意義

トラフィック特性としての自己相似性を検証する研究において，異常トラフィックによる非定常性が自己相似性の様相変化に影響すると指摘されながら，その性質を積極的に異常検知に応用するという検討はほとんど行われてこなかった．一方，本論文では，自己相似性評価手法の R/S 解析法から非定常性を定量化し，新たなトラフィック異常検知法を提案することを目的として検討を行った．以下に，本論文における工学的意義についてまとめる．

(1) R/S Pox Diagram において非定常的時系列を用いた検証により，R/S 解析法には定常的性質としての自己相似性を表現するだけでなく，トラフィック事象変化による非定常性を表現できる性質を有することを確認できた．上部プロット点群によりパケットトラフィック量の変化を表し，プロット点群の傾きが水平になることによりパケットが時間間隔をおいて到着する周期性を表すこ

Akita University

とを明らかにした．さらに，R/S統計量がある値に集約するときの任意長区間の区間長で周期を表すことを明らかにした．ここから，定常的性質と非定常的性質を同時に解析評価できる手法の確立が期待できる．

(2) R/S Poxレッグライン特性の周期推定による異常検知法は，従来法では検知しづらい低レートな長期的ポートスキャン攻撃を検知可能となることを明らかにした．ポートスキャンによる調査を把握可能となるため，ネットワーク管理においてセキュリティー対策に寄与できると期待される．

(3) 本研究では，TCP SYNパケットを観測対象としてポートスキャントラフィックの検知へ適用を検討したが，対象とする時系列を変更することでそこに現れる非定常性の性質も変わることが予想される．つまり，レベルシフトや周期性に基づく時系列解析として，様々な分野に対する応用も期待される．

5.3 今後に残された課題

最後に，今後に残された課題について述べる．

本研究では，R/S Poxレッグライン特性のSlope値導出範囲を一意に定めたため，あらかじめ周期推定範囲を想定した手法となっている．しかし，実トラフィックにおいては推定範囲外の周期的時系列も発生することも予想されるため，検討が必要である．この問題に対しては，導出範囲の調整や異なる周期推定範囲を定めた周期推定処理をマルチスレッドによって並列処理を実施することで解決すると考える．よって，実用化に向けたシステム開発の検討が今後の課題となる．

レベルシフトや周期性という非定常性に対する検討を行ったが，その他にも短時間に大量のパケットが到着する単パルス的トラフィックや，ワームやウィルス感染のように徐々にパケット数が増加するような非線形増加トラフィックなどの非定常性も観測される．この非定常性に対する提案特性の性能評価も検討する必要がある．

提案特性の特徴量で ST_{Inf} や SS_{Inf} という下限点群のSlope値が長期的ポートスキャンに対する検知法には有効性を示さなかったが，特異な変動傾向が観測される場合もあるため，その要因を検証することで，提案特性のさらなる拡張が期待できる．

Akita University

謝辞

本研究の遂行ならびに本論文の作成にあたって，終止懇切なるご指導とご鞭撻を賜りました秋田大学教授 工学博士 五十嵐隆治 先生 に心からお礼申し上げます．

本論文をまとめるにあたり，広い視野から数々の有益なご教示を頂きました秋田大学教授 工学博士 西田 眞 先生，東北公益文化大学公益学部長 工学博士 玉本 英夫 先生，秋田大学教授 博士（工学） 景山 陽一 先生，並びに同教授 博士（工学） 水戸部 一孝 先生に深く感謝いたします．

本研究の遂行にあたり，有益なご助言とご教示を賜りました東北大学教授 木下 哲男 先生，東北学院大学教授 岩谷 幸雄 先生，京都大学准教授 上田 浩 先生に深く感謝いたします．

本研究は，秋田大学 大学院工学資源学研究科情報工学専攻・工学資源学部情報工学科 五十嵐研究室において行われたものです．本研究の遂行において適切な助言を与えて下さった五十嵐研究室の皆様、卒業生の皆様に心から感謝いたします．

大学院後期課程への社会人入学について，ご配慮を頂きました秋田大学大学院工学資源学研究科情報工学専攻の関係各位に深く感謝いたします．

最後に，在学期間中，心身の支えとなってくれた家族に心から感謝いたします．

Akita University

発表論文

学術論文

1　レフェリー制のある学術雑誌

[1] 高橋秋典，五十嵐隆治，上田浩，岩谷幸雄，木下哲男，”R/S Pox レッグライン特性”，情報処理学会論文誌，54 巻，6 号，1761 頁～1770 頁（2013）

[2] A. Takahashi, R. Igarashi, H. Ueda, Y. Iwaya and T. Knottier, ”Network Anomaly Detection Based on R/S Pox Diagram”, International Journal of the Society of Materials Engineering for Resources, Vol. 17, No. 2, pp. 186-192 (2010)

[3] R. Igarashi, A. Takahashi, H. Ueda, Y. Nasuno, Y. Iwaya, M. Sakata and T. Kinoshita, ”A Proposal for Real Time Hurst Parameter Derivation”, Trans. IEE of Japan, Vol.127-C, No.6, pp.968-969 (2007)

[4] 瀧森徹, 高橋秋典, ”砂時計型ニューラルネットワークによる多価関数の近似”, 電気学会論文誌，第 120-c 巻 2 号，300 頁～301 頁（2000）

[5] 佐藤和人，菅原一隆，高橋秋典，瀧森徹，成田裕一，苗村育郎，”前頭葉萎縮の重症度とフラクタル次元の推定”，医療情報学，15 巻，4 号，247 頁～256 頁（1996）

[6] 佐藤和人，高橋秋典，瀧森徹，成田裕一，苗村育郎，”前頭葉萎縮に関する画像診断のための領域分割法”，医療情報学，15 巻，4 号，207 頁～216 頁（1996）

国際会議

[1] A. Takahashi, R. Igarashi, H. Ueda, Y. Iwaya and T. Kinoshita, ”Network Anomaly Detection Based on R/S Pox Diagram”, Proceeding of International Conference on Materials Engineering for Resources 2009, pp. 328-333 (2009)

[2] R.Igarashi, S. Ono, H. Inoue, A. Takahashi, T. Iwaya and M. Sakata, ”Some Features of Network Traffic Depending on Protocols”, Proc. of the 5th Inter. Conf. on Materials Engineering for Resources 2005 AKITA, BP-18, pp. 426-431 (2005)

[3] A. Takahashi, S. Abe, I. Namura and M. Nishida, ”Area Extract Method for Image

Diagnosis on Temporal Lobe Atrophy”，Proc. of the 3rd Inter. Conf. on Materials Engineering for Resources 1998 AKITA，BP-102，pp.101-102 (1998)

口頭発表

[1] 高橋秋典，五十嵐隆治，”情報セキュリティーポリシーを考慮したトラフィックデータ統計情報提供システムの試作”，平成25年度電気関係学会東北支部連合大会講演論文集（DVD-ROM），1F08（2013）

[2] 杉澤知，五十嵐隆治，高橋秋典，上田浩，岩谷幸雄，木下哲男，奈須野裕，”フロー量閾値設定のトラフィック特性の同定に関する研究”，平成25年度電気関係学会東北支部連合大会講演論文集（DVD-ROM），1F03（2013）

[3] 藤井俊，五十嵐隆治，高橋秋典，”フロー量閾値設定のトラフィック特性の同定に関する研究”，平成25年度電気関係学会東北支部連合大会講演論文集（DVD-ROM），1F03（2013）

[4] 加賀谷享諒，高橋秋典，五十嵐隆治，上田浩，岩谷幸雄，木下哲男，”R/S Pox レッグライン特性を用いたトラフィック状態判別法に関する研究”，第75回情報処理学会全国大会論文集（DVD-ROM），3Z-2（2013）

[5] 小西航，高橋秋典，五十嵐隆治，上田浩，岩谷幸雄，木下哲男，”ネットワークトラフィック変化検知のための視覚的表現法に関する研究”，平成24年度 第1回情報処理学会東北支部研究会講演資料（2012）

[6] 中尾拓也，高橋秋典，五十嵐隆治，上田浩，岩谷幸雄，木下哲男，”長期的ポートスキャントラフィックのパターン解析に関する研究”，平成24年度 第1回情報処理学会東北支部研究会講演資料（2012）

[7] 高橋秋典，五十嵐隆治，上田浩，岩谷幸雄，木下哲男，”R/S Pox レッグライン特性”，第11回情報科学技術フォーラム講演論文集，No.4，pp.9-16（2012）

[8] 中尾拓也，高橋秋典，五十嵐隆治，上田浩，岩谷幸雄，木下哲男，”長期的ポートスキャントラフィックのパターン解析に関する研究”，平成24年度電気関係学会東北支部連合大会講演論文集（2012）

[9] 小西航，高橋秋典，五十嵐隆治，上田浩，岩谷幸雄，木下哲男，”Pox Diagram 特徴量空間を用いたトラフィック変化検知”，平成24年度電気関係学会東北支部連合大会講演論文集（2012）

[10] 杉澤知，五十嵐隆治，高橋秋典，上田浩，岩谷幸雄，木下哲男，奈須野裕，”フロー量閾値設定のトラフィック特性の同定に関する研究”，平成24年度電気関係学会東北支部連合大会講演論文集（2012）

[11] 小西航，高橋秋典，五十嵐隆治，上田浩，岩谷幸雄，木下哲男，”ネットワー

Akita University

クトラフィック変化検知のための視覚的表現法に関する検討”，情報処理学会第57回CSEC・第17回IOT合同研究発表会 研究報告，Vol.2012-IOT-17，No.1，pp.1-6（2012）

[12] 高橋宏幸,高橋秋典,五十嵐隆治,上田浩,岩谷幸雄,木下哲男,奈須野裕,”ON/OFFモデルに基づくバックグラウンドトラフィック生成法に関する研究”，平成23年度 第2回情報処理学会東北支部研究会講演資料（2011）

[13] 鬼沢彩人，高橋秋典，五十嵐隆治，上田浩，岩谷幸雄，木下哲男，奈須野裕，”統計的な変化点検出法によるトラフィック異常検知”，平成23年度 第2回情報処理学会東北支部研究会講演資料（2011）

[14] 鬼沢彩人，高橋秋典，五十嵐隆治，上田浩，岩谷幸雄，木下哲男，奈須野裕，”統計的な変化点検出法によるトラフィック異常検知”，平成23年度電気関係学会東北支部連合大会講演論文集，pp.86（2011）

[15] 高橋宏幸,高橋秋典,五十嵐隆治,上田浩,岩谷幸雄,木下哲男,奈須野裕,”ON/OFFモデルに基づくバックグラウンドトラフィック生成法に関する研究”，平成23年度電気関係学会東北支部連合大会講演論文集，pp.85（2011）

[16] 中尾拓也，高橋秋典，五十嵐隆治，上田浩，岩谷幸雄，奈須野裕，木下哲男，”仮想マシンを用いたシミュレーショントラフィック生成に関する研究”，平成23年度電気関係学会東北支部連合大会講演論文集，pp.84（2011）

[17] 小西航，高橋秋典，五十嵐隆治，上田浩，岩谷幸雄，木下哲男，奈須野裕，”長期的スキャン攻撃の周期性に着目した異常検知法に関する研究”，平成23年度電気関係学会東北支部連合大会講演論文集，pp.83（2011）

[18] 高橋秋典，五十嵐隆治，上田浩，岩谷幸雄，木下哲男，”パルス型DoS攻撃におけるトラフィック特性の変化”，平成21年度電気関係学会東北支部連合大会論文集，pp.70（2009）

[19] 菊地征太郎,五十嵐隆治,高橋秋典,岩谷幸雄,木下哲男,奈須野裕,上田浩,”ns-2による実トラフィックモデル構成法の検討”，平成20年度第1回情報処理学会東北支部研究会講演資料，No.15（2008）

[20] 入江若菜，五十嵐隆治，高橋秋典，岩谷幸雄，木下哲男，上田浩，”Pox Diagram及び分散時間グラフに着目したトラフィック特性変化の研究”，平成20年度第1回情報処理学会東北支部研究会講演資料，No.14（2008）

[21] 高橋秋典，五十嵐隆治，上田浩，岩谷幸雄，木下哲男，”R/S Pox Diagramに基づくトラフィック異常検知に関する研究”，電子情報通信学会技術研究報告，Vol.108，No.203，pp.45-50（2008）

[22] 菊地征太郎,五十嵐隆治,高橋秋典,岩谷幸雄,木下哲男,奈須野裕,上田浩,”ns-2によるVoIPトラフィック疎通時のネットワーク特性の評価と検討”，平成20年

Akita University

度電気関係学会東北支部連合大会講演論文集，pp.103（2008）
[23] 入江若菜，五十嵐隆治，高橋秋典，岩谷幸雄，木下哲男，上田浩，”Pox Diagram 散布形状に着目したトラフィック状態推定法の検討”，平成20年度電気関係学会東北支部連合大会講演論文集，pp.102（2008）
[24] 高橋秋典，五十嵐隆治，入江若菜，上田浩，岩谷幸雄，木下哲男，”Pox Diagram 散布形状に基づくトラフィック異常検知法の検討”，平成20年度電気関係学会東北支部連合大会講演論文集，pp.101（2008）
[25] 五十嵐隆治，入江若菜，高橋秋典，正木忠良，佐々木芳宏，上田浩，奈須野裕，岩谷幸雄，木下哲男，”長時間トラフィックに含まれるトレンドの効果”，情報処理学会研究報告，Vol.2008，No.72，pp.47-51（2008）
[26] 相河浩之，高橋秋典，五十嵐隆治，”疎通プロトコルとネットワーク規模に依存するトラフィック特性に関する研究”，平成19年度第1回情報処理学会東北支部研究会講演資料，No.14（2007）
[27] 合田徹，高橋秋典，五十嵐隆治，岩谷幸雄，木下哲男，上田浩，奈須野裕，”ネットワークシミュレータを用いたVoIPトラフィック特性の検討”，平成19年度第1回情報処理学会東北支部研究会講演資料，No.15（2007）
[28] 斉藤楽，寺尾修二，五十嵐隆治，高橋秋典，玉本英夫，”シミュレータNS-2を用いたスループット特性の研究”，平成19年度第1回情報処理学会東北支部研究会講演資料，No.16（2007）
[29] 長谷川浩章，高橋秋典，五十嵐隆治，”R/S Pox-Diagramの特徴とトラフィックパラメータに関する研究”，平成19年度第1回情報処理学会東北支部研究会講演資料，No.17（2007）
[30] 高橋秋典，五十嵐隆治，上田浩，奈須野裕，岩谷幸雄，木下哲男，”オンラインネットワーク監視によるトラフィック異常検知”，電子情報通信学会技術研究報告，Vol.107，No.221，pp.57-62（2007）
[31] 合田徹，五十嵐隆治，高橋秋典，岩谷幸雄，木下哲男，上田浩，奈須野裕，”ネットワークシミュレータを用いたVoIPトラフィック特性の検討”，平成19年度電気関係学会東北支部連合大会講演論文集，pp.118（2007）
[32] 菊地征太郎，五十嵐隆治，高橋秋典，岩谷幸雄，木下哲男，上田浩，奈須野裕，”NS-2によるVoIPトラフィック疎通時のネットワーク特性の評価と検討”，平成19年度電気関係学会東北支部連合大会講演論文集，pp.117（2007）
[33] 高橋秋典，五十嵐隆治，岩谷幸雄，木下哲男，”R/S Pox Diagramにおけるプロット形状に着目したトラフィック特性の解析”，平成19年度電気関係学会東北支部連合大会講演論文集，pp.114（2007）
[34] 入江若菜，五十嵐隆治，高橋秋典，岩谷幸雄，木下哲男，上田浩，”非定常性

Akita University

を付加したシミュレーショントラフィックにおけるハーストパラメータの変化”，平成19年度電気関係学会東北支部連合大会講演論文集，pp.113（2007）

[35] 斉藤楽，寺尾修二，五十嵐隆治，高橋秋典，玉本英夫，”ネットワークシミュレータNS-2を用いたスループット特性の研究”，平成19年度電気関係学会東北支部連合大会講演論文集，pp.112（2007）

[36] 福士貴子，五十嵐隆治，高橋秋典，”モンテカルロ法によるトラフィックのモデリングの検討”，平成18年度第1回情報処理学会東北支部研究会講演資料，No.10（2006）

[37] 小野仁子，五十嵐隆治，高橋秋典，岩谷幸雄，坂田真人，”プロトコルに依存したネットワークトラフィックの解析”，平成18年度第1回情報処理学会東北支部研究会講演資料，No.9（2006）

[38] 合田徹，五十嵐隆治，高橋秋典，岩谷幸雄，上田浩，木下哲男，奈須野裕，”ネットワークシミュレータを用いたVoIPトラフィック特性の検討”，平成18年度電気関係学会東北支部連合大会講演論文集，pp.243（2006）

[39] 福士貴子，五十嵐隆治，高橋秋典，”モンテカルロ法によるトラフィックのモデリングの検討”，平成18年度電気関係学会東北支部連合大会講演論文集，pp.251（2006）

[40] 高橋秋典，五十嵐隆治，橋本芳明，岩谷幸雄，”実時間計測におけるハーストパラメータ高速導出法の検討”，平成18年度電気関係学会東北支部連合大会講演論文集，pp.250（2006）

[41] 小野仁子，田村雄介，増田有悟，五十嵐隆治，高橋秋典，岩谷幸雄，坂田真人，”プロトコルに依存したネットワークトラフィックの解析”，平成18年度電気関係学会東北支部連合大会講演論文集，pp.249（2006）

[42] 相河浩之，五十嵐隆治，高橋秋典，”侵入検知の可視化に関する研究”，平成18年度電気関係学会東北支部連合大会講演論文集，pp.246（2006）

[43] 長谷川浩章，五十嵐隆治，高橋秋典，”R/S Pox-Diagramの特徴とトラフィックパラメータに関する研究”，平成18年度電気関係学会東北支部連合大会講演論文集，pp.252（2006）

[44] 斉藤楽，五十嵐隆治，高橋秋典，坂田真人，”トラフィック時系列のフラクタル次元に関する研究”，平成18年度電気関係学会東北支部連合大会講演論文集，pp.253（2006）

[45] 田中政利，高橋秋典，五十嵐隆治，”照明条件の変化を考慮した肌色分布推定法”，第230回計測自動制御学会東北支部研究集会講演資料（2006）

[46] 田中政利，高橋秋典，五十嵐隆治，”背景関心領域から得られる指標を用いた動的肌色抽出法”，平成17年度第1回情報処理学会東北支部研究会講演資料

Akita University

（2005）

[47] 増田有悟, 五十嵐隆治, 高橋秋典, ”長期間観測に基づくネットワークトラフィック特性の検討”，成17年度第1回情報処理学会東北支部研究会講演資料，No.6（2005）

[48] 小野仁子, 田村雄介, 増田有悟, 五十嵐隆治, 高橋秋典, 岩谷幸雄, 坂田真人, ”ハーストパラメータとネットワークトラフィック量変化との相関”，平成17年度電気関係学会東北支部連合大会講演論文集，pp.162（2005）

[49] 小松武司，高橋秋典，田中政利，五十嵐隆治，”背景関心領域から得られる指標を用いた動的肌色抽出法”，平成16年度第1回情報処理学会東北支部研究会講演資料，No.10（2004）

[50] 田村雄介, 五十嵐隆治, 増田有悟, 高橋秋典, ”インターネットの長期トラフィックの性質に関する研究”，平成16年度第1回情報処理学会東北支部研究会講演資料，No.11（2004）

[51] 古戸寿幸，五十嵐隆治，田中翔，高橋秋典，”変動不感時間を付与できるランダムパルサーシステムの構築”，平成16年度第1回情報処理学会東北支部研究会講演資料，No.21（2004）

[52] 五十嵐隆治, 井上博勝, 宮林尚英, 田村雄介, 増田有悟, 岩谷幸雄, 坂田真人, 横山洋之, 藤原克哉, 高橋秋典, 玉本英夫, 行松健一, ”ハーストパラメータとネットワークトラフィック量変化との相関”, 情報処理学会研究報告, DSM-35, pp.7-12（2004）

[53] 田中政利，高橋秋典，小松武司，五十嵐隆治，”照度を用いた動的肌色領域抽出法に関する研究”，平成16年度電気関係学会東北支部連合大会講演論文集，pp.292（2004）

[54] 増田有悟, 五十嵐隆治, 高橋秋典, 玉本英夫, 横山洋之, 藤原克哉, 岩谷幸雄, ”分散－時間法、ペリオドグラム法におけるハーストパラメータ導出範囲の検討”，平成16年度電気関係学会東北支部連合大会講演論文集，pp.211（2004）

[55] 高橋秋典，小松武司，五十嵐隆治，”照度条件の変化を考慮した肌色領域抽出法の検討”，平成15年度電気学会C部門大会講演論文集，pp.912-913（2003）

[56] 田村雄介, 宮林尚英, 五十嵐隆治, 高橋秋典, 横山洋之, 玉本英夫, 岩谷幸雄, ”ハーストパラメータによるネットワークトラヒックの評価”，平成15年度電気学会C部門大会講演論文集，pp.867-873（2003）

[57] 宮林尚英，五十嵐隆治，高橋秋典，”ハーストパラメータによるネットワークトラヒック特性の解析”，平成14年度第1回情報処理学会東北支部研究会講演資料，No.5（2002）